DOCUMENT RESUME

ED 169 492 CS 004 753

TITLE	Evaluation of Outcomes, 1976-77: An Evaluation System Report on Reading Programs and Reading Achievement; Part IIA Technical Report.
INSTITUTION	District of Columbia Public Schools, Washington, D.C. Dept. of Research and Evaluation.
PUB DATE	Mar 79
NOTE	203p.; For related document, see CS 004 752; A number of pages in the appendixes may not reproduce well due to small or light type
EDRS PRICE	MF01/PC09 Plus Postage.
DESCRIPTORS	*Data Collection; Elementary Education; Program Evaluation; *Reading Achievement; *Reading Improvement; *Reading Programs; Reading Research; *Reading Skills; *Student Evaluation
IDENTIFIERS	District of Columbia

ABSTRACT

This document, prepared by the Evaluation System of the Public Schools of the District of Columbia, is the technical report of the results of analyses of factors affecting reading achievement for elementary students for the year 1976-77. The chapters include a discussion of the purpose and scope of the study; the methodology used; characteristics of the 1976-77 classes; results based on regression analyses, evaluations of reading programs in each grade level; and conclusions concerning reading achievement and testing, interest in reading, and evaluation of selected types of reading programs. Included are replicas of the classroom/student background form and the reading program form. (TJ)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED169492

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

Public Schools of the District of Columbia

EVALUATION OF OUTCOMES, 1976-77: AN EVALUATION SYSTEM REPORT ON READING PROGRAMS AND READING ACHIEVEMENT

PART II - A TECHNICAL REPORT

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Public Schools of the
District of Columbia
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

C S80 4 753

By
Division of Research and Evaluation
March 1979

Public Schools of the District of Columbia

OFFICE OF SUPERINTENDENT

Superintendent of Schools	Vincent E. Reed
Vice Superintendent of Schools	Elizabeth C. Yancey

DIVISION OF RESEARCH AND EVALUATION

Assistant Superintendent	Mildred P. Cooper
Director, Evaluation System	Herman Cobb, Jr.
Assistant Director	Beulah Glenn

Data Collectors (1976-77)

Dorothy Anderson	Isaac Jamison
Madeline Baker	Warren Kullberg
Linda Clifford	Carlos Reeder
Constance Dickerson	William Seldon
Veta Harrison	Olivia Simons

ACKNOWLEDGEMENTS

The Division of Research and Evaluation extends appreciation to the many individuals who have contributed both in general and in terms of this report to the Evaluation System of the Public Schools of the District of Columbia. Deep gratitude is expressed to the teachers, reading specialists, students, administrators, and others in the Evaluation System schools for their continuing cooperation and interest. We trust that they will find in the two parts of the report (the summary and the technical report) an indication of the kinds of information and evaluations that their efforts have made and will continue to make possible. We also thank the Division of Automated Information Systems (DAIS) which has handled the preparation and merging of computer tapes and those consultants--Dr. H. Russell Cort, Jr., Dr. Barry Smith, Dr. Walton Smith, and Bruce Jonas who provided major assistance in the preparation of this report.

PREFACE

This report is the second in a series of three Evaluation System reports to be issued by the Division of Research and Evaluation (DRE) by the end of the school year 1978-79.

This report provides an analysis of data from school year 1976-77. Report three will treat in a similar manner the data from school year 1977-78.

The first report presented an analysis of data from school year 1975-76.

Each report is in two parts: "Part I - A Summary" and "Part II - A Technical Report."

Additional reports will be developed and issued upon completion of further analysis of the data.

TABLE OF CONTENTS

		Page
LIST OF TABLES		VIII
LIST OF FIGURES		IX
CHAPTER		
I.	INTRODUCTION	1
II.	PURPOSE AND SCOPE	3
	Objectives	3
	The Data Collection and the Data	3
	Limitations	6
	Definitions	7
III.	METHODOLOGY	9
	Data Collection	9
	The Data Bank	9
	Future Developments in the Evaluation System	10
IV.	CHARACTERISTICS OF THE 1976-77 CLASSES	13
	Purpose of This Chapter	13
	The Variables Included	13
	Class-level Descriptive Statistics	15
	Characteristics of Reading Programs	32
V.	RESULTS BASED ON REGRESSION ANALYSES	42
	The Regression Analyses	42
	Regression on Posttest Performance Scores	49
	Regression Analysis of Reading Interest	63
VI.	EVALUATIONS OF READING PROGRAMS	79
	Methods	79
	Comparison of the First Grade Reading Programs	81
	Comparisons of Reading Programs at the Second Grade Level	87
	Comparisons of Reading Programs at the Third Grade Level	94
	Comparisons of Reading Programs at the Fourth Grade Level	105
	Comparisons of Reading Programs at the Fifth Grade Level	112
	Comparisons of Reading Programs at the Sixth Grade Level	117
	General	124
	Summary	127

TABLE OF CONTENTS (continued)

VII. CONCLUSIONS..... 135

Reading Achievement and Testing..... 135
Interest in Reading..... 136
Evaluation of Selected Types of Reading Programs..... 137

APPENDICES

A. REPLICA OF THE CLASSROOM/STUDENT BACKGROUND FORM..... 140

B. REPLICA OF THE READING PROGRAM FORM..... 143

C. EVALUATION SYSTEM SCHOOLS..... 148

D. CORRELATION MATRIX ACROSS ALL GRADES REGRESSION ON POSTTEST SCORES: Table A..... 150

E. CORRELATION MATRIX FOR PRIMARY GRADES: Table B..... 153

F. CORRELATION MATRIX FOR INTERMEDIATE GRADES: Table C..... 156

G. NUMBER OF STUDENTS AND CLASSES IN 34 EVALUATION SYSTEM SCHOOLS BY YEAR AND GRADE: Table D..... 159

LIST OF TABLES

Table		Page
I.	Characteristics of the 1976-77 Evaluation System Classes	16
II.	Benchmark Characteristics of Classes by Grade Level	29
III.	Percentages of Students in Different Reading Programs, by Grade and Year	35
IV.	Posttest Regression Summary: All Grades	51
V.	Posttest Score Regression-All Grades: Contribution of Sets	52
VI.	Posttest Regression Summary: Primary Grades	56
VII.	Posttest Score Regression - Primary Grades: Contribution of Sets	57
VIII.	Posttest Regression Summary: Intermediate Grades	60
IX.	Posttest Score Regression - Intermediate Grades: Contribution of Sets	61
X.	Reading Interest Regression Summary: All Grades	65
XI.	Reading Interest Score Regression - All Grades: Contribution of Sets	66
XII.	Reading Interest Regression Summary: Primary Grades	70
XIII.	Reading Interest Score Regression - Primary Grades: Contribution of Sets	71
XIV.	Reading Interest Regression Summary: Intermediate Grades	75
XV.	Reading Interest Score Regression - Intermediate Grades: Contribution of Sets	76
XVI.	Statistical Data for Each Reading Program Compared, Grade 1	83
XVII.	Comparison of Reading Programs by Receiving Extra Help, Grade 1	84
XVIII.	Composition of Samples for Analyses of Reading Programs, Grade 1	85
XIX.	Statistical Data for Each Reading Program Compared, Grade 2	89
XX.	Comparison of Reading Programs by Receiving Extra Help, Grade 2	91

LIST OF TABLES (Continued)

Table		Page
XXI.	Composition of Samples for Analyses of Reading Programs, Grade 2	93
XXII.	Statistical Data for Each Reading Program Compared, Grade 3	95
XXIII.	Comparison of Reading Programs by Receiving Extra Help, Grade 3	96
XXIV.	Compositions of Samples for Analyses of Reading Programs, Grade 3	97
XXV.	Statistical Data for Each Reading Program Compared, Grade 4	107
XXVI.	Comparison of Reading Programs by Receiving Extra Help, Grade 4	108
XXVII.	Composition of Samples for Analyses of Reading Programs, Grade 4	109
XXVIII	Statistical Data for Each Reading Program Compared, Grade 5	113
XXIX.	Comparison of Reading Programs by Receiving Extra Help, Grade 5	114
XXX.	Composition of Samples for Analyses of Reading Programs, Grade 5	115
XXXI.	Statistical Data for Each Reading Program Compared, Grade 6	118
XXXII.	Comparison of Reading Programs by Receiving Extra Help, Grade 6	120
XXXIII.	Composition of Samples for Analyses of Reading Programs, Grade 6	121
XXXIV.	Types of Reading Programs Evaluated, 1975-76 and 1976-77, by Grade	129

LIST OF FIGURES

Figure		Page
1.	Regions of Significant Difference Between Pairs of Reading Programs, Grade 3	99
2.	Regions of Significant Difference Between Pairs of Reading Programs, Grade 4	110
3.	Regions of Significant Difference Between Pairs of Reading Programs, Grade 6	122

I. INTRODUCTION

The Evaluation System of the Public Schools of the District of Columbia has several purposes. At the most general level it is designed to enable the school system to evaluate the effectiveness of programs and to identify characteristics of programs that seem to increase effectiveness. Also, it is designed to enable the D.C. Public Schools to follow particular students over a period of years, to examine long-term effects of different programs, and to measure changes in program characteristics from year to year. The design of the Evaluation System is such that analyses or evaluations can be done at a number of levels: the individual student, the classroom, the school building, the grade level, the program, the region, and the overall school system.

An essential feature of the Evaluation System is its ability to correlate data. Thus, the System is designed to combine test data with program data, and to combine both of these with a variety of other data or measures such as: student absence, sex, age, size of class, student turnover rate for classes, type of class, involvement of aides, and many other such measures. These data or measures come from various sources. Combining them is a major undertaking. This report represents the second year of analysis of various student, class, teacher, and program data combined with performance measures based on test data.

The Evaluation System is designed to support the investigation of an almost limitless number of questions. It is thus the intention of the Division of Research and Evaluation, which developed and implements the

Evaluation System, that users of its reports (administrators, teachers, and others) become familiar with the System and its potential. The objective findings of the Evaluation System will provide decision-makers at all levels with sound information upon which to make judgments for future directions.

II. PURPOSE AND SCOPE

A. Objectives

This is the second report of this series. It presents results of analyses of factors affecting reading achievement as measured by the percentage of objectives mastered on the Prescriptive Reading Test (PRT), and by students' interest in reading for the year 1976-77. It also presents descriptive information about students, classes, and reading programs for that year. Data for grades 1-6 are included in the report. The report also provides comparisons with results obtained from analyses of the 1975-76 data. Conclusions and recommendations based on the two years of data and analyses are presented. In some cases, identified trends lead to clear conclusions. In other cases, changes from one year (1975-76) to the next (1976-77) are noted and discussed and in such cases it is noted that analyses of the 1977-78 data will usually be needed to identify reliable trends or patterns.

B. The Data Collection and the Data

1. Collection Procedures

The data in this report was collected using three basic instruments.

a. The Classroom/Student Background Form (SBF) is a two-part machine-readable form designed especially for the Evaluation System by the Division of Research and Evaluation (DRE). One part of the form is for the collection of background data on the classroom and the other part is for the collection of background data on the students in the class. This form is completed by staff from DRE in the fall and spring of the school year. See Appendix B.

b. The Prescriptive Reading Test (PRT), the school system's own criterion-referenced reading test, is administered in the fall and spring of each year. The test is constructed by levels ranging from A through E and G which cover the reading grade level objectives for grades 1-6. Students are administered test levels according to their instructional level. Fall and spring test results were matched for each student by DRE.

c. The Reading Program Form (RP) is a two-part machine-readable form designed by DRE and completed by teachers during the month of May of each year. One part of the form is for the collection of classroom reading data and the other part is for the collection of individual student reading program data. See Appendix C.

Data was collected from the 34 elementary schools participating in the Evaluation System. The official October 1976 total membership in these schools for grades 1-6 was 15,382 students. See Appendix A for the distribution of schools by region.

2. Test Data

The test data are percentages of reading objectives mastered in the fall of 1976 (pretest) and in the spring of 1977 (posttest) on the Prescriptive Reading Test (PRT), Levels A through E and G. Only the scores for students who had both pre and post scores for the same level of test were used. Since many teachers administered different levels of the PRT to different students in their classes (as they should have), test raw scores were converted to percentages in order to arrive at a class average, or in order to make the range and units of test scores reasonably

comparable for different students in the same reading program. The "raw score" for a test was the number of objectives mastered according to PRT criteria. The percentage computed was the number of objectives mastered divided by the total number of objectives in the test for the level administered. The number of objectives varied from 21 for Level A to 45 for Level B. Subsequent reports will examine subtest scores, computed in similar fashion.

3. Student Characteristics

Data, other than test scores, which pertain to individual students include sex, age, absences, total time enrolled in a class, student interest in reading, need for extra help in reading, how often student reads on own and student's class and ID number.

4. Teacher Characteristics

Characteristics of teachers include absences, reasons for using particular reading materials, freedom to choose own reading skills objectives to be taught, extent to which teachers coordinated instructional reading program with other school personnel and length of time, and period of time during the year the teacher was associated with a particular class.

5. Reading Programs

A description of each student's reading program in terms of materials, grade levels of materials, and relative emphasis given to each material was obtained from the reading form described in Item 1.c. While the reading program is a characteristic of each student, it is listed separately here as a reminder that a focus of the Evaluation System is programs. The programs, of course, may differ even for students within the same classroom.

6. Classroom Characteristics

Division of Research and Evaluation staff collected data pertaining to the classroom as a whole, not specifically to individual students (e.g., number of students eligible for free lunch program, number of students referred for various reasons, type of teaching situation, presence of an aide, availability of reading or mathematics specialists, and the like). In addition, a number of measures pertaining to the classroom as an aggregate were computed from data based on individuals (e.g., student absence rate, student turnover rate, class average percentage of objectives mastered, class average interest in reading, and class average need for extra help in reading).

C. Limitations

It is important that the reader understand the limitations on the generalizability of the results and conclusions reported here. First, there were large losses for some variables due to missing data. For example, on test data given to the Division of Research and Evaluation the pretest or posttest (or both) data were missing from the records of about 31% of the 15,382 students. Second, for purposes of this report, only test data for students who were administered the same level of test, pre and post, were used. That automatically excluded all students who were administered a higher or lower level test in the fall, (pretest) than in the spring, (posttest), although those students will be included in subsequent analyses. The total number of matched test level students was 6,008, or 39% of the total population. Third, there were losses of data for particular variables owing to the procedures used for computer tape files. Fourth, there were losses of data for particular variables owing to the failure of some forms to be filled out accurately or

completely, but also owing to errors in optical scan processing of data from forms. The limitations described are mainly the consequences of continued limitations of resources, processing and the like, not of design limitations of the Evaluation System.

In a data gathering program of this scope, it was infeasible to evaluate the consistency of ratings of students made by teachers. That is, it was not feasible to do a form of test-retest. By the same token, it was not feasible to observe reading programs in the classrooms. All such reading data are based on the subjective reports of teachers.

Finally, as will be noted, it is important to temper certain conclusions based on analyses of the 1975-76 and 1976-77 data until cross-validation can be made with data from 1977-78. The 1977-78 data analyses will be provided in the third report of this series by June 1979.

The findings reported in the Summary (Part I) are stated in non-technical terms without detailed statistical back-up. Detailed statistical analyses, as well as technical discussions, are found in this Technical Report (Part II) upon which the Summary Report was based.

D. Definitions

Classroom Characteristics - Include grade, room number, number of free lunch pupils, class size, and percent of male and female.

Learning Barriers - Those conditions which may confront a student and hinder the learning process, i.e., poor attendance, lack of proper nutrition, economic conditions, and medical/physical health.

Posttest Scores - The percentage of reading objectives mastered in the spring based on the total number of objectives possible at the test level administered.

PMT - Prescriptive Mathematics Test

Pretest Scores - The percentage of reading objectives mastered in the fall based on the total number of objectives possible at the test level administered.

PRT - Prescriptive Reading Test

Reading Behaviors - Those behaviors a student exhibits toward reading, i.e., amount of interest in reading and the degree to which student reads on his own.

Teacher Characteristics - Those descriptors of the teacher such as the day the teacher started with the class, the last day with the class, and teacher turnover rate.

III. METHODOLOGY

A. Data Collection

Since the present report is intended to parallel the earlier technical report of the 1975-76 data, the basic approach to data collection, briefly described in Chapter II, was the same as for the earlier year.

Data were collected in the form of reading pretest and posttest scores, and test results for each student are available both in a composite score and objective by objective. In addition, the forms completed by teachers and Division of Research and Evaluation staff members provided a variety of important information concerning individual students and the classroom in which they were taught. The specific variables and sets of variables on which the data were collected are described in more detail in Chapter IV.

Once the forms had been completed, they were edited, scanned, and stored on magnetic tape. The test scores were also provided for both fall and spring and stored on tape.

B. The Data Bank

In the earlier technical report, we described the development of the Evaluation Data Acquisition, Control, and Analysis (EDACA) System as a component of the Evaluation System. As was noted in the earlier report, we are dealing, even in the pilot study, with data from some 15,382 students forming 551 classes in grades 1-6 in 34 schools. A computer-based data acquisition, storage, and processing system like EDACA is essential to the handling of such large quantities of data.

Using both the 1975-76 data base and the 1976-77 data base in part for the purpose of testing the EDACA System, it has been possible to make considerable progress on that system. The evolving concept is that of an efficient

computer-based system that will be highly standardized, yet flexible enough to permit the incorporation and deletion of variables and the addition of larger numbers of schools, classes, and students, eventually to include the entire school system if desirable.

Computer Programs. The EDACA system involves a number of different computer programs, each fulfilling a specific need. Basically, however, there are three types of programs involved in the system: (a) manipulation programs permit the programmer to store and access the data files as well as to merge two or more files into a single file and to create new files containing partial or specialized information needed for particular analyses; (b) descriptive statistical programs provide information concerning frequencies, means (averages), and other basic statistics which help us to describe and thus better understand the nature of the classes and students appearing in our samples. These programs thus tell us, for example, how many classes from the total sample of classes appear in each of the six grades, how many classes have experienced a change of teacher during the school year, the average class size, the average number of class absences, and the average percentage of male students in the classes (out put from these descriptive programs is dealt with in some detail in Chapter IV); and (c) analytical programs permit various statistical analyses of the data to be undertaken. The two principal types of analysis applied to both the 1975-76 data and the 1976-77 data included in the present report are regression analysis and analysis of covariance. These are described in two later chapters.

C. Future Developments in the Evaluation System

The EDACA System has been further developed beyond the point at which it stood when the first study was completed. It has now been partially standardized to provide the kinds of file manipulation, statistical description,

and analysis included in both the present and earlier reports. The system is now being expanded to readily permit further types of analysis and thus permit us to learn even more from the data bases provided by the evaluation system as it currently exists. Plans for the further development and modification of the EDACA System still involve the same major goals as stated in the earlier technical report:

1. Flexibility

The system is now flexible enough that it will permit reasonable additions and deletions of variables as new questions are raised and old ones are satisfactorily answered. Further work remains to fully document the flexibility of the system and to further develop that flexibility.

2. Expansion Capability

Although the evaluation system does currently involve a large number of students, classes, and schools, it is recognized that this still represents a pilot stage in its development. Future plans call for the extension of the system to a larger number of elementary schools, to secondary schools, and eventually to the entire school system. This means that the capability for expansion in numbers of schools, classes, and students must be built into the developing EDACA system. It now seems clear that quantitative expansion of the computer analysis system will be straightforward. However, it is essential to recognize that some kinds of expansion may necessitate the continued use of a large scale computer system, such as that currently being employed.

3. Reduction of Data Loss

Data loss has been considerably reduced in the present study as compared with the earlier study of 1975-76 data. Through refinement of data processing procedures, it was possible to identify 396 classes valid for analysis, compared to the 229 classes identified in the earlier year. Further work is being done to produce an even greater reduction in data loss in future years.

4. Individualized Feedback

The present report focuses on data descriptions and analyses which cut across all available schools, classes, and students. This level of data processing is essential in providing important information concerning adjustments and modifications that can be made by the school system in its ongoing efforts to improve the reading and other skills of the individual child. It would also be desirable, however, to provide direct feedback to teachers, principals, and regional and central administrative personnel concerning their specific classes, schools, or regions. Thus, it would be helpful to the individual teacher to know in detail the characteristics of his or her classroom and the individual students in the classroom and to know how these characteristics compare with those generally characterizing similar classrooms in the school system. Such individualized profiles are not yet available, but it is hoped that the EDACA system can be expanded to include such a profiling procedure.

5. Rapid Turnaround

An important goal in the further development of the EDACA system is to provide rapid processing of data, such that feedback can be provided within a relatively short time following the collection of data. For example, it would be desirable to provide feedback early in the fall semester concerning data collected the previous spring. This has become one of the major goals of the Evaluation System, and EDACA represents, in part, an effort to achieve this goal. Over the course of the two studies completed to date, turnaround time has been considerably reduced, and it is expected that gradual further reduction in turnaround time is a reasonable goal for the future.

IV. CHARACTERISTICS OF THE 1976-77 CLASSES

A. Purpose of This Chapter

This chapter presents descriptive statistics pertaining to the 1976-77 Evaluation System schools. The data presented are for the samples used in major analyses (Chapters V & VI), with a few exceptions to be noted in the text. Descriptive statistics are presented for both classes and students as the units of analysis. Thus, the reader will find that there may be differences in the values of a particular variable for those two conditions. It is always necessary to keep in mind whether one is considering classes or students when considering the data.

B. The Variables Included

The descriptive statistics to be presented here are for variables that were used in analyses of relationships and differences, as noted above. For class-level fixed-order regression analyses of achievement, a total of 22 variables (21 predictors and 1 criterion variable) were used, including pretest and post-test measures. Regression analyses of interest in Reading used one additional predictor. Exploratory regression analyses employed 34 predictor variables, however, thirteen were eliminated from the achievement analyses reported in Chapter V. The aim was generally to eliminate those variables which did not appear to show a significant and consistent relationship to the achievement criterion variable, and thus to maximize useable sample sizes.[1/]

For student-level analyses of covariance, 2 covariates were selected. These were variables for which the measures pertained to the individual student.

1/ Given sets of analyses always used a list-wise deletion procedure to determine samples. Any case for which any variable was missing was thus excluded. Under that procedure losses of cases can rise rapidly as more variables are included if there are missing data. Thus, it is important to minimize the number of variables used.

To maintain sample sizes of classes in regression analyses, regressions were made for primary level samples of classes (grades 1-3) and for intermediate level samples of classes (grades 4-6). In effect, primary grade classes were analyzed as a group, and intermediate grade classes were analyzed separately as another group. Thus, the descriptive statistics presented in this chapter pertain to these two main groups of classes as well as to the total group (grades 1-6). Multi-grade classes were included in these groups by taking the unweighted average grade of each multi-grade class. Thus, a 4-5 combination class would have a computed grade level of 4.5. A 4-5-6 multi-grade class would be assigned a grade level of 5. With the exception of certain analyses (which will be noted), class-level statistics are for sets as described, not for each grade level separately. Given this method of assigning grade levels to individual classes, the actual upper boundary of primary classes was 3.4. The lower boundary of intermediate classes was 3.5. The boundaries of all classes were, of course, 1-6.

Analyses of covariance of reading programs were done on a grade by grade basis. The analyses of covariance used the student as the unit of analysis. In one set of analyses, only students from single grade level classes were included. In a second set of analyses, students from multi-grade classes were included. The basis for assigning students from multi-grade classes to a particular grade level (1, 2, 3...6) for purposes of analysis was to define the lower and the upper boundary as -.5 and +.4 respectively. Thus, a student from a multi-grade class with an assigned grade level of 3.4 would be included in an analysis of covariance of 3rd grade students. That would be the case even if the student literally was in the 4th grade part of the class. The reason for this is that grade level code pertains to the class in the Evaluation System, not to the student within the class.

C. Class-level Descriptive Statistics

The total analysis pool of classes for the 34 elementary schools in the Evaluation System in 1976-77 was 551.[2/] That is an average of 16 classes per school, or an average of 92 per grade level if only 6 grades are considered. It is an increase of 39 classes over the 1975-76 pool. The maximum useable samples were:

. For all grades combined: 396 classes

. For primary grades: 191

. For intermediate grades: 205

Thus, the useable sample of 396 classes was about 72% of the total pool.[3/]

Table I gives descriptive statistics for Primary grades (1-3), Intermediate Grades (4-6), and the Total Useable Sample for all grades (1-6). The data for the three samples are shown in the same table to make it easy to compare the three groups on each variable. One should remember, however, that Primary classes actually range from grades 1-3.4, Intermediate ones from grades 3.5-6. Definitions of variables follow in the text below. The reader is reminded that the data in Table I refer to classes as units, not to individual students as units.

1. Definitions of Variables in Table I

Variables in Table I are defined below. The variables are lettered here as they are in Table I to aid the reader.

a. Pretest % Objectives Mastered. This is the average of the class percentages of objectives mastered on the Prescriptive Reading Test (Fall

2/ 551 was not the total number of classes in the Evaluation System. It was the total number meeting certain ID matching criteria. The total number of classes in the Evaluation System was 570 for 1976-77.

3/ The total analysis pool of 551 classes was reduced to 396 useable classes through a list-wise deletion process. That is, if any student in a class was missing data on any one of the selected variables the whole class was lost. Only students with pre-post matched data on the same level PRT were used, although class level measures such as class size, turnover rate, % males, etc., were computed from total determined class membership.

Table 1

Characteristics of the 1976-77 Evaluation System Classes

Name of Variable	Level*	# of Classes	Mean	SD	Correlation With: Posttest**	Interest
a. Pretest % Objectives Mastered	Primary	191	58.06	11.77	.68	-.28
	Inter.	205	58.27	11.58	.75	-.20
	Total	396	58.27	11.66	.68	-.24
b. Posttest % Objectives Mastered	Primary	191	77.19	11.73	---	-.40
	Inter.	205	69.82	10.62	---	-.24
	Total	396	73.38	11.75	---	-.34
c. Grade (Ave. Grade Level)	Primary	191	1.67	.99	.29	-.19
	Inter.	205	4.96	.86	-.08	.01
	Total	396	3.38	1.89	-.22	.03
d. Test Level Count (Ave. # Different Test Levels Class)	Primary	191	1.83	1.02	-.15	-.00
	Inter.	205	2.37	1.03	.19	.10
	Total	396	2.11	1.06	-.08	.07
e. Class Size	Primary	191	25.5	4.80	.11	.03
	Inter.	205	26.93	5.54	-.01	-.02
	Total	396	26.28	5.23	.00	.01
f. % Males (Class %)	Primary	191	50.02	10.47	-.12	.24
	Inter	205	50.27	11.17	-.04	.12
	Total	396	50.18	10.82	-.08	.18

* Note: Primary = Primary Grades 1 through 3.4.
Inter. = Intermediate Grades 3.5 through 6.
Total = Total Useable Sample

**Correlations of the variable designated with posttest are specifically with arcsine transformations of posttest class average percentages of objectives mastered.

Table 1 (Continued)

Name of Variable	Level*	# of Classes	Mean	SD	Correlation With: Posttest**	Interest
g. Class Holding Rate (100-Turnover Rate)	Primary	191	87.82	9.87	.01	-.11
	Inter.	205	90.02	8.17	.18	-.13
	Total	396	88.96	8.57	.04	-.11
h. Teacher Turnover	Primary	191	.03	.17	-.05	-.00
	Inter.	205	.03	.17	.02	-.08
	Total	396	.03	.17	-.01	-.04
i. Teacher Absence Proportion	Primary	191	.05	.04	-.06	.00
	Inter.	205	.06	.06	-.02	.06
	Total	396	.05	.05	-.06	.04
j. Proportion Free Lunch	Primary	191	.51	.39	-.35	.11
	Inter.	205	.51	.33	-.27	.25
	Total	396	.51	.35	-.29	.18
k. Amount Reading on Own	Primary	191	1.69	.28	-.45	.76
	Inter.	205	1.78	.28	-.17	.75
	Total	396	1.74	.28	-.34	.76
l. Average Interest in Reading	Primary	191	2.00	.45	-.40	---
	Inter.	205	2.08	.42	-.24	---
	Total	396	2.04	.44	-.34	---
m. Ave. Need for Extra Help Rdg.	Primary	191	2.14	.41	.41	-.39
	Inter.	205	2.19	.39	.27	-.42
	Total	396	2.16	.40	.31	-.40
n. % Not Receiving Help Elsewhere	Primary	191	39.05	31.87	.08	.07
	Inter.	205	47.68	33.27	.11	-.01
	Total	396	43.52	32.85	.04	.04

Table 1 (Continued)

Name of Variable	Level*	# of Classes	Mean	SD	Correlation With: Posttest**	Interest
o. Learning Problems-Reading	Primary	191	1.94	1.39	-.05	-.06
	Inter.	205	1.69	1.27	-.13	.09
	Total	396	1.81	1.33	-.05	.00
p. Learning Barriers-Behavior	Primary	191	.88	.91	-.18	.13
	Inter.	205	1.01	1.01	-.13	.14
	Total	396	.95	.96	-.17	.14
q. Learning Barriers-Medical Physical Health (Vision, Sickness, etc.)	Primary	191	.64	.96	-.01	-.09
	Inter.	205	.40	.81	-.01	.09
	Total	396	.51	.89	.03	-.02
r. Average Level of 1st Rdg Material	Primary	191	5.57	1.98	.20	-.18
	Inter.	205	9.60	1.50	.08	-.25
	Total	396	7.66	2.67	-.15	-.06
s. SD of Level of 1st Rdg Material	Primary	191	1.64	.82	-.18	.16
	Inter.	205	1.51	.74	-.23	.32
	Total	396	1.58	.78	-.16	.23
t. Presence of Math Specialist	Primary	191	.82	.39	.22	-.21
	Inter.	205	.69	.46	.14	-.12
	Total	396	.75	.43	.12	-.14
u. Teaching Situation-Proportion Open Space	Primary	191	.31	.47	-.34	.14
	Inter.	205	.20	.40	-.21	.10
	Total	396	.26	.44	-.23	.11
v. Teaching Situation-Proportion Team Tchg	Primary	191	.36	.48	-.28	.21
	Inter.	205	.26	.44	-.07	.12
	Total	396	.31	.46	-.14	.16

* * * * * *

Table 1 (Continued)

Name of Variable	Level*	# of Classes	Mean	SD	Correlation With: Posttest**	Interest
The following variables were included in some regression analyses. Note the change in sample sizes.						
aa. Mean Class Absences	Total	354	11.61	4.67	-.10	.05
bb. SD of Test Level	Total	354	.48	.43	-.17	.17
cc. Representativeness of Reading Program (%)	Total	354	98.89	6.47	.06	-.01
dd. Mean Duration of Students in Rdg Class (Ave. Scale Value)	Total	354	7.21	.91	.17	-.19
ee. Class Average Need for Extra Help in Reading	Total	354	2.16	.40	.33	-.39
ff. Presence of Aide	Total	354	.29	.46	.06	.00
gg. Presence of Reading Specialist	Total	354	.83	.38	.00	-.08
hh. Teaching Situation Departmentalized	Total	354	.08	.28	.12	-.14
ii. Teaching Situation Multi-Age Class	Total	354	.23	.42	-.03	.06
jj. Teaching Situation Multi-Level Class	Total	354	.25	.43	-.02	.05
kk. Taught by Other Teacher	Total	354	.28	.45	-.08	-.00
ll. Learning Problems-Math	Total	354	1.34	1.48	.12	-.11
mm. Learning Barriers-Economic Conditions (Clothing, Shoes, etc.)	Total	354	.28	.53	.03	.04

1976). The number of objectives mastered at a given test level was converted to the percentage of total objectives possible for that level. This was done for each student with matched pre-post test data for the same test level. The percentages for all matched levels in a given class were combined to give a weighted average for the class. The mean percentages in Table 1 are the unweighted average of those weighted averages for the designated groups. The correlation shown for this variable with posttest is the correlation of the arcsine transformation of pretest with the arcsine transformation of posttest. A class is a homeroom class, since that is how students are grouped for testing.

b. Posttest % Objectives Mastered. Defined same as Pretest % Objectives Mastered, except pertains to posttest (Spring 1977). This is one of two main out-come measures used in the studies reported here.

c. Grade. This variable is the grade level of the class. The means in Table 1 are the average grade levels of the groups designated. The grade level of multi-grade classes was obtained by taking the unweighted average of the grades included in a given class. A weighted average would have been obtained by taking into consideration the number of students at each grade level within the class.

d. Test Level Count. This is the number of levels of PRT administered pre and post (matched in a class.) The means in Table 1 are the unweighted averages for the designated groups. The measure for a given class does not take into account the number of students receiving each test level.

e. Class Size. This is the total number of students who were recorded in the Evaluation System as members of a given class according to certain inclusion rules. A student was considered a member of a class if the data for the student could be matched with a valid teacher and school ID number,

regardless of whether or not the student had pre or posttest data, and if the student was a member of the class for 100 days or more (i.e., End Day minus Start Day = 100 or greater). Thus, "class size", as the term is used here, is not the actual class size from the teacher's point of view, but the collectivity of students meeting certain criteria for purposes of analysis. In the sample of 369 classes for which data are presented in Table 1, the average class size (as we have defined size) was 26. The reader needs to keep in mind that the term "class size" has a special meaning in this report.

f. % Males. This is the number of males divided by the number of members of a class (as defined in Class Size) times 100. It is the complement of the percentage of females.

g. Class Holding Rate. This is the complement of the class turnover rate. The turnover rate was obtained essentially by dividing the number of students who were not members of a class for the whole school year by the total number of members, and multiplying by 100. That number was then subtracted from 100.

h. Teacher Turnover Proportion. Any class that had a teacher who did not remain with a class for the school year was coded 1; all others were coded 0. The means in Table 1 are the proportions of classes in the designated groups that had a turnover of teachers by that definition.

i. Teacher Absence Rate. This is the ratio of the number of days the teacher was absent to the total number of school days counted for the year for the teacher. For 1976-77 the Evaluation System terminated data collection on school day #162; that is the maximum number of teacher days possible for purposes of computing this measure.

j. Proportion Free Lunch. This is the number of students in a class reported by a teacher to be eligible for the free lunch program divided by

the aggregate number of students constituting a class as defined in Class Size. This measure was intended to provide an index of the socioeconomic composition of each class. It is important to recognize that the basic datum in this index is the number designated by the teacher for the whole class at a given point in time, not a designation for each individual student.

k. Amount of Reading on Own. This is the average for the class of the teacher's ratings of each student on the following scale:

"Aside from reading instruction, how often does this student read on his own in school? (library books, comics, etc.)"

1 = Often

2 = Sometimes

3 = Never

Ratings were made in April 1977.

l. Average Interest in Reading. This is the average rating for the class made of each student by the teacher on the following scale:

"How would you assess this child's interest?"

1 = Very interested

2 = Moderately interested

3 = Seldom interested

4 = Not interested

5 = Actively dislikes reading

This measure, like other ratings of individual students by their teachers, was made in the late Spring of 1977. Note in Table I that the scale values, as with those for the preceding measure, are such that the more positive the rating, the lower the scale value. Thus the correlations of this measure with Posttest % Objectives are consistently negative; the correlations of this variable with Amount of Reading on Own are consistently positive (and high). One needs to bear in mind the direction of scales when one looks at the

correlations in Table I in order to state the relationships correctly. In Chapter V, correlations involving these measures have had their signs reversed to aid the reader in making more direct interpretations.

This is the second main outcome measure used in the analyses reported here.

m. Average Need for Extra Help in Reading. This is the average, for the class, of ratings made of each student by the teacher on the following scale:

"Is this child in need of extra help in reading?"

1 = Yes, very much

2 = Yes, somewhat

3 = No

As with all such measures, the means, standard deviations and correlations of this measure are based on the distributions of it in the group designated in Table I.

n. % Receiving Help Elsewhere. This is the number of students in a class rated by the teacher as receiving extra reading instruction during the regular school day from one or more sources other than the teacher, divided by the number of students rated; the ratio was multiplied by 100 to make it a percentage.

o. Learning Problems-Reading. This is the number of students in a class rated by the teacher as having been referred for reading difficulties on the following scale:

1 = 1-3 Students

2 = 4-6 Students

3 = 7-9 Students

4 = 10 or more Students

Note that the ratings for the measure are strictly scale values; the measure does not come from ratings of individual students.

p. Learning Barriers-Behavior. This measure is based on the same scale and procedure as Learning Problems-Reading.

q. Learning Barriers-Medical/Physical Health. This measure is based on the same scale and procedure as Learning Problems-Reading.

r. Average Level of First Reading Material. This is a rating made for each student by the teacher of the level of the reading instruction material used most frequently with the student during the month preceding the rating, which occurred in the later spring of 1977. Level was based on the following scale:

1 = Reading readiness

2 = Pre-primer

3 = Primer

4 = 1^1 (grade one, first semester)

5 = 1^2 (grade one, second semester)

6 = 2^1

7 = 2^2

8 = 3^1

9 = 3^2

10 = 4

11 = 5

12 = 6

13 = 7

14 = 8

15 = 9

s. SD of Level of First Reading Material. This is the standard deviation of the levels of the most frequently used reading instruction materials for the class. It was included in analyses as a measure of the heterogeneity of reading instruction for each class. The means and standard deviations of this measure in Table I pertain to the distributions of class standard deviations in the designated groups of classes.

t. Presence of Math Specialist. The school indicated whether there was a math specialist available to the class by checking "Yes" or "No".

u. Teaching Situation-Proportion Open Space. The school indicated whether or not the class was open-space or self-contained. The means in Table I are the proportions of open-space classes in the groups designated.

v. Teaching Situation-Proportion Team Teaching. Same code values and procedure as for Teaching Situation-Proportion Open Space. These variables are not mutually exclusive. The teacher could indicate the presence or absence of up to 5 teaching situation variables.

With the exception of variable aa below, the remaining variables were included in exploratory regression analyses, but were not included in the set of variables used in the regression analyses based on classes as the units in Chapter V of this report. They are, nevertheless, descriptive of classes in a total useable sample of similar size to that finally used. Thus, they are included in the final part of Table I.

aa. Mean Class Absences. This is the average number of days that students in a class were reported to have been absent during the 1976-77 school year for the Evaluation System schools. The average pertains to a maximum of the first 162 school days of that year. It was obtained by dividing the total number of days absent of all students in the class by the total number of days all students were members of the class (up to a limit of 162 days).

bb. SD of Test Level. The standard deviation of test levels for each class. Levels of the PRT were numbered (A = 1, B = 2, C = 3 ...) and treated as if they represented points on a continuum. The standard deviation for each class was investigated as a measure of class diversity of reading levels.

cc. Representativeness of Reading Program. Teachers were asked to rate each student's reading program on the following scale:

> "To what extent is the reading program described above indicative of what was used all year for this child for reading instruction? (Exclude changes in reading level.)"
>
> 1 = Don't know-student new to this class
>
> 2 = Very representative or very similar
>
> 3 = Somewhat representative or somewhat similar
>
> 4 = Not at all representative or not at all similar

A measure of representativeness for each class was obtained by dividing the total number of 2's and 3's by the total number of 2's, 3's and 4's and multiplying by 100.

dd. Mean Duration of Students in Reading Class. (Average Scale Value.) On the Reading Program Form (DCDRE #200) for each student the teacher rated how long the student had been in the class that year on the following scale:

1 = less than one month

2 = 1 month

3 = 2 months

4 = 3 months

5 = 4 months

6 = 5 months

7 = 6 months

8 = 7 months

9 = 8 months

The measure for each class is the average of ratings of time in class for the class.

ee. Class Average Need for Extra Help in Reading. This is the average for the class of the rating for each student on the following scale:

"Is this child in need of extra help in reading?"

1 = Yes, very much

2 = Yes, somewhat

3 = No

ff. Presence of Aide. Each school indicated on a form whether or not the class had an aide.

gg. Presence of a Reading Specialist. Each school indicated whether or not a reading specialist was available to the class.

hh. Teaching Situation-Departmentalized. Same procedure as for variables u and v.

ii. Teaching situation - Multi-Age Class. Same procedure as for variables u and v.

jj. Teaching Situation - Multi-Level Class. Same procedure as for variables u and v.

kk. Taught by other Teacher. Teachers indicated the number of their homeroom students who were given regular reading instruction by some other teachers.

ll. Learning Problems - Math. Same procedure as for variable o.

mm. Learning Barriers - Economic conditions. (clothing, shoes, etc..) Same procedure as for variable p and q.

2. Benchmark Characteristics of Classes by Grade Level.

In order to provide benchmarks against which to compare achievement measures for years following 1976-77, selected statistics were computed for

classes in each of the six basic grade levels in the Evaluation System. These benchmark statistics are shown in Table II.

The classes at each grade level in Table II came from the 551 classes described as the total useable sample. Only classes that were single grade level classes were included. That is, multi-level classes were not included. The variables in Table II are the same as those defined above for Table I. Therefore, no further descriptions of these variables is needed for Table II. It is, perhaps, appropriate to remind the reader that classroom averages for pretest and posttest percentage of objectives mastered are composites of the percentages for the different test levels administered pre and post. As Table II shows for the variable Test Level Count, there was typically more than one PRT test level used in the classes at each grade level. Indeed, the average number of different test levels used for all of the 376 classes in Table II was 2.0.

In passing, there are two very striking trends worth noting in Table II. The correlation of pretest and posttest increases consistently from Primary to Intermediate grades. Also, Posttest % Objectives drop about 10% from Primary to Intermediate groups of classes. These trends also occurred in the 1975-76 data.

There are changes in the benchmark characteristics from the 1975-76 to the 1976-77 pools of data. The major difference is the increase in average class sizes. That change, however, is attributable to improvements in the basic data files and to a change in the method and criteria used for aggregating. Despite those changes, the other variables shown in Table II are very similar for the two years.

Of particular interest are the average pre-post differences in % of objectives mastered for the six grade levels for the two years:

Table 11

Benchmark Characteristics of Classes by Grade Level

Characteristics	Grade Level 1	2	3	4	5	6	1-6
Number of Classes	57	62	53	65	63	76	376
Pretest % Objectives Mast'd	58.96	57.24	63.82	57.07	59.93	58.17	59.04
SD	10.75	11.30	13.16	11.29	10.22	13.15	
r_{xy}†	.53*	.64*	.65*	.79*	.71*	.79*	
Posttest % Objectives Mast'd	78.66	78.61	79.92	69.80	69.13	68.48	73.64
SD	11.11	9.95	11.83	9.39	10.71	12.73	
Pre-Post Difference	19.70	21.37	16.10	12.73	9.2	10.31	14.60
Class Size	25.60	26.15	24.64	26.80	25.73	26.77	26.02
SD	4.50	5.49	4.35	6.18	4.29	5.82	
r_{xy}	.26*	.12	-.13	.04	-.05	.13	
Mean Class Absences	12.77	11.35	10.73	11.86	11.36	11.67	11.63
SD	3.46	3.87	4.11	4.53	4.43	4.47	
r_{xy}	.09	.08	-.19	-.14	-.16	-.12	
% Males	51.22	50.48	53.15	49.00	51.40	49.24	50.62
SD	8.13	11.05	11.41	9.92	9.33	12.87	
r_{xy}	-.02	-.03	-.22	.00	-.32*	-.14	
Class Holding Rate	87.42	88.37	84.46	88.87	89.59	90.81	88.46
SD	9.09	9.11	10.59	8.94	8.21	7.65	
r_{xy}	-.04	-.01	.14	.01	-.02	.38*	
Test Level Count	1.37	1.60	1.75	2.23	2.35	2.51	2.00
SD	.67	.71	.76	.90	1.02	1.14	
r_{xy}	-.02	.15	.06	.29*	.21	.39*	

* $p \leq .05$

† r_{xy} is the correlation of the variable with Posttest % Objectives Mastered (without the arcsine transformation of Posttest).

Note: Data are for students in single grade classes only. Multi-grade and combination grade classes are not included.

	Grade						
	1	2	3	4	5	6	Wtd. Ave.
1975-76	25.23%	19.95%	12.48%	12.62%	9.90%	5.44%	14.00%
1976-77	19.70	21.37	16.10	12.73	9.20	10.31	14.60
Difference:	-5.53	+1.42	+3.62	+ .11	-.70	+4.87	+ .60

The 3rd and 6th grades show a noticeable increase, while the 1st grade shows a decrease. The average change for the six grades was a slight increase. A two-way analysis of covariance (ANCOVA) was performed to test the null hypothesis that the difference from 1975-76 to 1976-77 was not significant. The independent variables were Year (1975-76, 1976-77), and Grade (1-6); the dependent variable was average percentage of posttest reading objectives mastered. The covariate was average percentage of pretest objectives mastered. The unit of analysis was the class. The summary table of the analysis was:

Source	df	F	p
Year (Y)	1	.619	.432
Grade (G)	5	62.893	.001
Y x G	5	3.752	.002
Regression	1	1279.919	.001
Error	922		

Thus, there was no overall significant difference between the two years (Y, as the +.60 difference would suggest in any case). The differences between grades (G) were significant, but of greater interest was the significant Year by Grade (Y x G) interaction. Since the interaction was highly significant, an F-test for each grade level was made of the 1975-76/1976-77 adjusted differences. The results were:

Statistic	Grade 1	2	3	4	5	6
F	3.310	0.107	3.825	0.167	0.293	14.154
df	1,163	1,139	1,152	1,171	1,159	1,132
p	.071	.744	.052	.684	.589	.001*+

* = $\leq$ .050
\+ = $\leq$ .0083, the Bonferroni p for a critical value of t ($F = t^2$) when 6 independent tests are made, assuming familywise alpha = .05; i.e., critical $p_B = .05/6 = .0083$.

It can be seen that by the usual .05 criterion, the change for grade 6 was significant. By a conservative Bonferroni criterion, the change for that grade was also significant. It may be concluded that there was a reliable increase in percentage of objectives mastered from pretest to posttest by 6th grade classes in 1976-77, compared to 1975-76.

What is the meaning of this finding? Something clearly seems to have changed at the 6th grade level. Whether the significant gain means more effective teaching at the 6th grade level in 1976-77 can not yet be concluded with assurance. There are alternative hypotheses that could explain the observed result. If it is found with the 1977-78, and the 1978-79 data that the same amount of gain or more occurs, there will then be unambiguous support for the hypothesis that instruction has improved. Such findings will be especially important for their implications for the effectiveness of the Competency-Based Curriculum (CBC) program.

Another comparison of the gains is of importance. Both in 1975-76 and 1976-77, the primary grades showed larger average gains than the intermediate grades:

	1975-76	1976-77
Primary Grades	19.22%	19.06%
Intermediate Grades	9.32%	11.37%
Difference	9.90%	7.69%

The difference between the two levels of grades, however, was less in 1976-77 than in 1975-76. It is clear that the decrease in average unweighted difference was the result of an increase in gain by the intermediate grades (4, 5, and 6) rather than a decrease by the primary grades (1, 2, and 3). If the primary grades are taken as a standard, then a very direct measure of continuing improvement in instruction in the intermediate grades will be a continuing decrease in the average difference in gains in reading objectives mastered between the two levels. A measure of improving instruction at both levels will be a continuing rise in average gain at the primary level in conjunction with even greater rises in gain at the intermediate level. There is, of course, a limit to the amount of gain both levels can show, since it is impossible to have more than 100% objectives mastered at posttest. Furthermore, the increasing gain measure is based on the assumption that pretest levels stay about the same from year to year, as they did in the results described here. It is also based on the assumption that the same tests continue to be used, with about the same mixes of test levels at each grade level, as indeed was the case in '75-'76 and '76-'77.

As always, of course, we need to bear in mind that these results, so far, pertain to the population of students that is administered the same level of PRT pre and post, and for which both pre and post test data are available. That population is about 70% of the total population of students in the Evaluation System for which pre-post test data are available to date. Subsequent Evaluation System reports will of course note any changes in the specification of populations to which results apply.

D. Characteristics of Reading Programs

Since Chapter VI presents comparisons of the effectiveness of selected reading programs, it is appropriate to provide descriptions of the relative prevalence of programs. For purposes of this report, a reading program is

defined as the types of reading instruction materials used most frequently, and next most frequently, with each student for the month preceding the completion of the Reading Program Form by the student's teacher near the end of the spring semester.

The Reading Program Form (see Appendix B) lists reading instruction material under seven major categories:

- Basal Series
- Linquistic/Phonic Materials
- Programmed/Structured Materials
- Language Experience Materials
- Individualized Reading Materials
- Supplementary Materials
- Special Materials

In describing a student's reading program, the teacher indicated the materials used most frequently, next most frequently, and third most frequently. For each material reported, the teacher also rated the reading level of each material (see definition of variable in Section C.1.r. above). Each material listed on the Reading Program Form falls uniquely into one of the seven main types or categories of materials listed above. Thus, one way of identifying a reading program is to consider only the main categories of materials listed for each student. As noted above, only first and second most frequent categories were used to define programs in reports of the '75-'76 and the '76-'77 data.

Table III shows the prevalence of different programs, as defined, for grades 1-6. This listing is for students only from single grade level classes. Prevalence here means the number of students classified as having had a particular program, regardless of the specific material or reading level of the material. Table III also shows the programs selected for comparison at

each grade level (see Chapter VI). In Table III the order in which categories are listed for a given program is: most frequently used, next most frequently used. Also, some programs appear to involve only one category of material; e.g. Basal, Linguistic/Phonic, and Language Experience. Such programs mean that the teachers indicated either just a single material, or two materials falling in the same category; for example, a basal text and nothing else; or two different basal texts.

It can be seen in Table III that the most prevalent (in terms of percentages of students) type of program is a basal and a linguistic/phonics combination.

The prevalence of different types of reading programs for 1975-76 and 1976-77 was generally similar. There were some changes, however, that may be indicative of changing trends in the utilization of different theories of reading instruction. There was a marked <u>increase</u> in the use of a pure linguistic/phonics type of reading program in first and third grade classes (from 2.9% in '75-'76 to 10.3% in '76-'77 for grade 1; 3.2% to 8.9% for grade 3). There was a marked <u>decrease</u> in the prevalence of basal-programmed/structured types of programs in grades 2 and 3 (13.2% to 4.9% grade 2; 11.8% to 1.3% in grade 3). Except for grade 6, there was a general increase in the prevalence of pure basal programs, a general decrease in prevalence of basal-supplementary materials types of program. The latter trend was especially pronounced in primary level classes.

If major categories of materials are considered individually, a comparison between years can be made of the relative prevalence of each category, regardless of whether it was the most frequently used or the next most frequently used. For all six grades combined, the changes were as follows:

Table III

Percentages of Students in Different Reading Programs, by Grade and Year

Program†	Year	Grade 1	2	3	4	5	6	1-6
Basal	'76-7	13.9*	14.4*	17.4*	14.0*	18.0*	9.8*	14.4
	'75-6	12.5*	7.1*	10.4*	7.7*	13.4*	11.8*	10.6
Basal Ling/Phon	'76-7	29.8*	34.0*	34.1*	34.5*	33.6*	30.9*	32.7
	'75-6	35.4*	36.0*	33.2*	42.9*	32.7*	34.8*	35.7
Basal-Programmed	'76-7	6.8	4.9*	1.3	1.2	1.6	0.8	
	'75-6	9.6*	13.2*	11.8*	2.4	0.6	0.7	6.0
Basal-Lang. Exp	'76-7	14.1*	10.3*	6.7*	3.2*	2.3	0.2	6.1
	'75-6	10.5*	3.4	4.1	4.5	4.1*	0.0	---
Basal-Ind. Rdg	'76-7	1.1	13.0*	11.7*	12.2*	12.1*	9.2*	11.4
	'75-6	0.8*	9.0	11.2*	9.9*	22.1*	9.2*	11.0
Basal-Suppl.	'76-7	4.6*	4.9*	7.3*	16.3*	6.5*	13.2*	8.9
	'75-6	11.7*	12.4*	12.3*	19.6*	9.1*	17.8*	13.7
Basal-Special	'76-7	0.4	0.1		0.3	0.1		
	'75-6	0.0	0.0		0.0	0.1	0.1	
Ling/Phon	'76-7	10.3*	6.2*	8.9*	6.2*	3.9*	7.5*	7.2
	'75-6	2.9	8.1*	3.2	6.2	1.6	7.6*	
Ling/Phon-Basal	'76-7	1.7	3.6*	1.6	0.1	1.8	1.1	
	'75-6	5.8	2.9	1.3	0.3	0.8	0.9	
Ling/Phon-Prog'd	'76-7	1.4	0.6	0.0		0.4	0.7	
	'75-6	2.7	1.0	0.2		0.1	0.7	

Table III (Continued)

Program†	Year	Grade 1	2	3	4	5	6	1-6
Ling/Phon-Lang. Exp.	'76-7	1.2	0.3	0.1	0.7	2.1	0.1	
	'75-6	0.1	0.0	0.0	0.0	0.7	0.0	
Ling/Phon-Ind. Rdg.	'76-7	2.6	1.5	2.5	4.9	3.8*	1.2	
	'75-6	0.0	0.2	3.7	4.2	3.5*	2.9	
Ling/Phon-Suppl.	'76-7	3.6*	0.1	1.8	1.4	0.1	3.8	
	'75-6	0.8	0.1	0.8	0.3	1.4	0.1	
Ling/Phon-Special	'76-7	0.8	0.1	0.1	0.6	0.4	0.4	
	'75-6	0.1	0.0	0.0	0.0	0.0	0.0	
Programmed/Struct'd	'76-7	0.9	2.0	0.8	0.2	0.3	0.2	
	'75-6	1.4	0.0	0.5	0.0	0.0	0.1	
Prog'd-Basal	'76-7	0.3	0.2	3.4				
	'75-6	1.7	5.5*	0.8				
Prog'd Ling/Phon	'76-7	0.3	2.0	0.2		0.1		
	'75-6	0.0	0.0	2.5		0.0		
Prog'd Lang. Exp.	'76-7	0.9	0.1	0.1		0.1		
	'75-6	1.1	0.0	0.2		0.0		
Prog'd-Ind. Rdg	'76-7		0.1	0.1	0.1	0.3	0.1	
	'75-6		0.0	3.4	0.3	0.1	1.0	
Prog'd-Suppl	'76-7	0.4	0.8		0.6	0.1		
	'75-6	1.0	0.0		0.0	0.0		
Prog'd-Special	'76-7	0.2		0.3	0.4		0.1	
	'75-6	0.0		0.0	0.0		8.0	

Table III (Continued)

Program†	Year	Grade 1	2	3	4	5	6	1-6
Language Experience	'76-7	0.6	0.2				0.2	
	'75-6	0.2	0.1		0.1		0.0	
Lang Exp-Basal	'76-7	1.0	0.0					
	'75-6	0.2	0.2	0.6	0.1			
Lang Exp-Ling/Phon	'76-7	0.1	0.1					
	'75-6	0.5	0.0					
Lang Exp-Prog'd	'76-7	0.1						
	'75-6	0.0						
Lang Exp-Ind. Rdg	'76-7	0.5	0.1			0.1	0.1	
	'75-6	0.0	0.0			0.1	0.2	
Lang Exp-Suppl.	'76-7	0.5						
	'75-6	0.4						
Lang Exp-Special	'76-7							
	'75-6							
Indiv'd Reading	'76-7	0.1		0.1	0.1	1.7	2.3	
	'75-6	0.0		0.0	0.1	0.0	0.3	
Ind. Rdg-Basal	'76-7	0.2	0.2	1.3	0.1	1.6	0.0	
	'75-6	0.0	0.0	0.0	0.0	0.1	0.1	
Ind. Rdg-Ling/Phon	'76-7					1.6	1.5	
	'75-6				0.8	2.9	0.1	
Ind. Rdg-Prog'd	'76-7	0.1						
	'75-6	0.0						

Table III (Continued)

Program†	Year	Grade 1	2	3	4	5	6	1-6
Ind. Rdg-Lang. Exp.	'76-7					2.3	0.9	
	'75-6					2.0	1.5	
Ind. Rdg-Suppl.	'76-7		0.2	0.1	1.2	2.9	0.9	
	'75-6		0.0	0.0	0.2	0.0	2.1	
Ind. Rdg-Special	'76-7							
	'75-6					0.1	0.1	
Supplementary	'76-7	1.2		0.1	0.1	0.1	5.6*	
	'75-6	0.1	0.1	0.0	0.2	1.3	0.8	
Suppl-Basal	'76-7	0.0		0.1	1.7		0.5	
	'75-6	0.2		0.0	0.0	0.3	0.0	
Suppl-Ling/Phon	'76-7	0.1				0.1	3.5	
	'75-6	0.0	0.6		0.3	0.6	3.9	
Suppl-Prog'd	'76-7							
	'75-6		0.1					
Suppl-Lang. Exp.	'76-7	0.1				0.2	0.2	
	'75-6	0.1		0.1		0.0	0.1	
Suppl-Ind. Rdg.	'76-7	0.0				1.8	0.8	
	'75-6	0.3				1.9	2.4	
Special	'76-7	0.3					0.9	
	'75-6	0.0				0.1	0.0	
Special-Basal	'76-7					0.1	1.6	
	'75-6					0.0	0.0	

Table III (Continued)

Program†	Year	Grade 1	2	3	4	5	6	1-6
Special-Ling/Phon	'76-7	0.2	0.1					
	'75-6	0.0	0.0	0.1				
Special-Prog'd	'76-7							
	'75-6			0.1				
Special-Ind. Rdg.	'76-7						0.1	
	'75-6						0.0	
Special-Lang. Exp	'76-7							
	'75-6					0.1		
Special-Suppl.	'76-7						0.1	
	'75-6						0.0	
Total No. Students	'76-7	1772	1433	1439	1634	1539	1794	9611
	'75-6	1066	979	1197	1189	1427	1067	6925

*Programs selected for comparison within a grade level (see Chapter VI).
†Abbreviations: Ling/Phon = Linguistic/Phonics
Programmed or Prog'd = Programmed/Structured
Lang. Exp - Language Experience
Ind. Rdg = Individualized Reading
Suppl = Supplementary

Type of Material	1975-76	1976-77	Difference
Basal	84.8%	79.5%	-5.3%
Linguistic/Phonics	48.5%	49.6%	+1.1%
Programmed/Structured	11.1%	5.9%	-5.2%
Language Experience	5.8%	8.3%	+2.5%
Supplementary	18.0%	14.9%	-3.1%
Special	1.4%	1.2%	-0.2%

From this it can be seen that there has been a general shift away from basals, programmed/structured and supplementary materials, and a shift toward using linguistic/phonic and language experience types of reading materials. The shifts, of course, were not uniform at each grade level. This can easily be seen by comparing the average unweighted differences for primary and intermediate grades:

Types of Material	Primary	Intermediate
Basal	-6.5%	-6.3%
Linguistic/Phonics	+2.6%	-1.1%
Programmed/Structured	-9.5%	-2.2%
Language Experience	+5.6%	-0.3%
Supplementary	-5.4%	-1.4%
Special	+0.7%	-1.4%

Thus, it can be seen that:

- there was little difference in the shift away from basals between the two levels;
- the shift toward using linguistic/phonics materials occurred predominantly in the primary grades;
- the overwhelming shift away from using programmed/structured materials occurred in the primary grades;
- overwhelmingly, the increase in use of language experience materials and supplementary materials was in the primary grades;

- the shift away from special materials occurred mainly in the intermediate grades.

These shifts in the prevalence of different types of reading materials used suggest important changes in the theories and methods of reading instruction occurring throughout the school system. The consequences of these changes will reveal themselves over time in the achievement of students. Analyses of the 1977-78 data will be made to determine whether these trends continue, and if they do, how they are affecting reading achievement.

V. RESULTS BASED ON REGRESSION ANALYSES

The earlier Evaluation System Report on the 1975-76 data contained a number of regression analyses. This section of the present report provides a parallel set of regressions on the 1976-77 data base. In what follows, we report first on the nature of regression analysis and the particular variables included in our analyses. This is followed by a section on the specific regression analyses carried out on the 1976-77 data. Results for the two years are treated comparatively at appropriate points.

The regression analyses were carried out in an effort to determine what variables, from within the set described above, are significantly associated with and prediction of reading performance and reading interest. The practical implications of such an approach to data analysis are straightforward. Basically, such analyses can help us to determine what prior test performance characteristics, classroom characteristics, teaching practices, and student characteristics actually "make a difference" in the sense of being statistically associated with reading performance and interest. Knowledge of which variables are predictive of performance and interest scores will provide for the school system an empirical basis for making possible modifications which can reasonably be expected to lead to increases in either reading performance or reading interest or both. It is important to keep in mind that the results reported in this chapter refer to classes, not to individual students.

A. The Regression Analyses

In order to accomplish our basic purpose of learning what sets of variables and what variables within those sets are significantly predictive of reading performance and interest scores, a series of regression analyses were performed. In each regression analysis, there is a single variable, termed

the criterion variable, which we are trying to predict and a number of different variables that we have reason to believe may be associated with the criterion. These latter variables are termed predictors and are examined, individually and in sets, to determine which ones do, in fact, significantly predict the criterion. To clarify the nature of the regression analysis, consider briefly the example of a baseball manager who wishes to increase the batting averages of his players. Taking a systematic regression approach, he might determine the existing batting averages of a large number of players and collect information concerning each player's height, weight, muscle volume, amount of experience, age, and the training methods employed in teaching that player to bat. He would enter all of these variables into a regression analysis to determine which would significantly predict the criterion "score," the player's batting average. With this approach, he can learn which variables are important in the training of players in the future and may thus learn, for example, that players with greater height, but not necessarily with greater weight, who have greater muscle volume and have been trained with certain training methods have higher batting averages.

1. The Criterion Variables

The "batting averages" of interest in the present study were the posttest reading performance score and the student's interest in reading as rated by the teacher. The reading performance score employed was basically the percentage of objectives mastered (Posttest % Objectives). This percentage was determined for each individual student, and the individual percentages were averaged to determine the average or mean score for each class. This percentage score, termed the posttest score, is described in more detail in Chapter IV.[4/]

4/ The posttest score was subjected to an arcsine square root transformation before regression analyses were carried out.

The second criterion score employed in separate regression analyses was the student's interest in reading. This score is based on an observer rating scale, completed by the student's teacher as an item on one of the forms described in Chapter II. For each student, the teacher was asked to assess the child's interest in reading by indicating which of the following best described the student's interest level:

1 = Very interested

2 = Moderately interested

3 = Seldom interested

4 = Not interested

5 = Actively dislikes reading

For statistical purposes, each of these terms represented a point on a 5-point rating scale, where a high score of 5 is assigned to "very interested" and a low score of 1 to "actively dislikes reading" (intentionally reversed).

2. Predictor Variables

The variables used in an attempt to predict the posttest performance and reading scores have been described. For purposes of analysis, the variables finally selected for entry were grouped into nine major sets. In fact, four of these "sets" consisted of a single variable, while the remaining 5 contained between 2 and 5 variables each. Following is a brief description of each set and a listing of the variables within that set. Sets and variables are described in the order in which they were entered into the hierarchical regression analyses. For a more detailed treatment of the individual variables and the distributional characteristics associated with these variables, the reader is referred back to Chapter III.

a. Set 1 - Pretest Score. Also a single-variable set, pretest score is the class average percentage of correct objectives and represents the fall

reading performance test score of the individual child averaged over the classes as a whole.[5/] The entry of the pretest score at an early point in the hierarchical regression is important, since it served to partial out, and thereby control, the effect of the student's reading level at the beginning of the year when examining his/her reading level at the end of the year as measured by the posttest score.

b. Set 2 - Grade. This set consists of a single variable, and is simply the grade of which the class is a part. Grades 1 through 6 are represented.

c. Set 3 - Test Level Count. This set also consists of a single variable, which is simply a count of the number of different test levels represented by children in a given class. It is, in effect, a measure of the degree of homogeneity of grouping, where homogeneity is defined in terms of test levels. A very homogeneous class might have only a single level of testing represented in the entire class (all students at the same level), while a relatively heterogeneous class might have several different levels represented.

d. Set 4 - Basic Classroom Characteristics. This fourth set consists of three variables, each contributing to a basic description of the class being entered into the analysis. In utilizing these variables, it was considered possible that such characteristics as class size might contribute significantly to the prediction of the posttest reading score or the reading interest score or both. The three variables in the set were:

- Class size
- Percent male (sex distribution)
- Class holding rate

5/ As with the posttest score described above as a criterion variable, the pretest score was subjected to an arcsine square root transformation before regression analyses were carried out.

e. Set 5 - Teacher Behavior Characteristics. This set of variables describes two potentially important aspects of teacher behavior, characteristics which we had reason to believe might be predictive of students' posttest scores. The two major variables included in the set are:

. Teacher turnover rate

. Teacher absence rate

f. Set 6 - Percent Free Lunch. This set, consisting of a single variable, indicates the percentage of students in a given class receiving free lunch. The variable may provide some indication of the socioeconomic level represented in the class and again may be predictive of reading interest or performance.

g. Set 7 - Reading Behavior Characteristics. Consisting of 4 major variables, the reading behavior set deals with certain student characteristics (averaged over the class as a whole) that are directly involved with reading and reading activities and may well be predictive of reading performance. The four variables are:

. Average amount of reading on own

. Average interest in reading

. Average amount of need for extra help in reading

. Average percentage receiving help elsewhere

h. Set 8 - Learning Problems and Barriers. Certain kinds of problems faced by the individual student may serve as barriers to effective learning. The inclusion of Set 8 in the analyses represents an attempt to examine the relationship of the prevalence of problems and barriers in a class to class performance or interest scores. The three major variables contained in the set are:

. Health barriers

. Reading problems

. Behavior barriers

i. Set 9 - Classroom Environment. The learning experiences of the individual student in the classroom setting are, to some degree, determined by certain kinds of variables which, together, comprise what may be termed the classroom environment. The present analyses entered several of these potentially important variables in an attempt to identify those aspects in the classroom environment which may, in fact, relate significantly to reading performance and interest. The five major variables involved in this final set are:

. Mean level of most-used reading material

. Presence of a math specialist

. Open space versus self-contained classroom

. Team teaching vs non-team teaching

. Variability (standard deviation) of level of most-used reading material

3. Specific Regression Analyses.

As with the 1975-76 data, a series of six regression analyses were carried out. The first three of these involved the class average posttest score as the criterion variable and the nine sets of variables noted above as predictors. The first analysis involved the total valid sample of 396 classes (except those missing data on specific variables) cutting across all six grades. It should be noted that this represents a substantial gain over the 229 classes available for analysis in the 1975-76 data. The larger sample size in the more recent data base reflects primarily continuing progress in the development and improvement of the Evaluation System. While it is important to know which variables and sets of variables significantly predict posttest across all grades, it was

considered likely that somewhat different variables might significantly predict the posttest scores at primary and intermediate grades. As a result, a second regression analysis employed the posttest scores and associated predictor scores in the first three grades (1, 2 and 3), while a third analysis examined relationships in the three intermediate grades (4, 5 and 6).

The remaining three analyses paralleled those described above, except that the criterion variable was the class average interest in reading rather than posttest reading performance score. In addition, for these analyses, it was necessary to remove reading interest and average amount of reading, which closely parallels the interest score, as predictors. Average class absence rate was added as a variable.

It should be noted that, while the regression analyses on the 1976-77 data largely parallel those carried out on the 1975-76 data, it was considered desirable to make one important change in the nature of the analyses. Specifically, the Grade set was the first set entered in the earlier data analyses, and the Pretest Score variable was entered third. This meant, statistically, that the effective pretest score was not partialled out or taken into account before the Grade and Test Level Count sets were examined. While the ordering decision must necessarily be a somewhat arbitrary one, it was considered desirable in the 1976-77 analyses to remove the Pretest Score effect first, rather than third. This assures that any differences in pretest scores are taken into account before the effects of any other variables, including grade and test level, are considered. The change in ordering may have affected the statistical impact of these three variables, but should not affect the impact of any of the remaining variables entered as later sets in the analysis.

4. Limitations

Before presenting the results of the regression analyses, it is important to recall that the EDACA system and the Evaluation System more generally

are still being developed and tested. While there have been definite improvements in both the EDACA System and the data base for the current analyses, the data base still necessarily has certain limitations. The sample of 396 total classes considered valid for analysis is certainly much larger than the sample of 229 classes available in the 1975-76 data. However, there is still considerable data loss, and the loss may be systematic, rather than totally random. Thus, the final sample in the present analyses may be a biased one and may not be a fully representative sample of the classes in the 34 pilot schools. The problem is not a serious one, since we will be cross-validating present findings, both by comparison with the 1975-76 data and through later examinations of the 1977-78 data and similar data for later years.

The results of each of the regression analyses on posttest performance scores, as well as those on reading interest scores, are provided in two separate sections. The first provides a brief statement of the statistical results of the analyses, while the second provides a discussion and interpretation of the results.

B. Regression on Posttest Performance Scores

As has been noted, three regression analyses involving the posttest reading scores as the criterion variable were conducted. The first cuts across all six grades, the second deals with the primary grades only, and the third deals with the intermediate grades only.

1. Means and Standard Deviations

Table I in Chapter IV provides the means and standard deviations for all variables across all grades and separately for the primary three and the intermediate three grades.

2. The All-Grades Regression

Results of the first regression analysis are dealt with in two separate sections, as noted above.

a. Results. Table A (in Appendix D) provides the matrix of correlations among the predictors and the criterion variable involved in the regression. Table IV indicates the multiple correlation (R), the R^2 value, the R^2 change, and the simple correlation associated with each variable in the fixed order regression. The final R and R^2 values indicate that the variables entered resulted in a significant correlation of .78, which accounts for approximately 61% of the variance. A more conservative estimate suggests a likely population R^2 of .60.

In order to provide an appropriately conservative estimate of variables contributing significantly to the regression equation, F-ratios were first calculated to determine whether or not a given set of variables made a significant contribution. Only if the set contribution was significant were the F-ratios for individual variables calculated. It should also be noted that, for both sets and individual variables, the F-ratios calculated were based on a Model II (Cohen & Cohen, 1975) approach, in which the error term involves all variables entering into the total analysis.

Table V provides the F-ratios and associated R^2 change values for the nine sets of variables entered into the regression. It is clear from the table that five of the nine sets significantly contributed to the prediction of posttest reading scores. The first set, Reading Pretest Score, was entered primarily as a control variable in order to partial out the effect of pretest performance on the further prediction of posttest scores. As would be expected, pretest is positively associated with posttest and makes a reliable contribution to the regression equation. The second set, Grade, was also significant and, like pretest score, was entered primarily in order to partial out the effect of the variable.

The third significant set, the percentage of students receiving free lunch, was negatively associated with posttest scores. This suggests that

Table IV

Posttest Regression Summary: All Grades

Variable	Multiple R	R Square	RSQ Change	Simple R	F-Ratio		B	Beta
Pretest Score	.67537	.45613	.45613	.67537	436.45***	†	.70393	.61053
Grade	.72965	.53239	.07627	-.22347	72.97***	†	-.01101	-.14817
Test Level Count	.73177	.53549	.00310	-.07653	SNS	†	.00327	.02481
Class Size	.73294	.53720	.00171	.00457	SNS		.00059	.02185
% Males	.73296	.53723	.00003	-.08028	SNS		.00018	.01357
Class Holding Rate	.73327	.53768	.00045	.03944	SNS		-.00019	-.01187
Teacher Turnover	.73484	.53998	.00230	-.01208	SNS		-.05658	-.06932
Teacher Absence	.73611	.54185	.00187	-.05808	SNS		-.08168	-.03041
% Free Lunch	.73970	.54716	.00531	-.29362	5.08*		-.02722	-.06824
Reading on Own	.75408	.56864	.02148	-.34196	20.55***	†	-.00999	-.02024
Reading Interest	.75781	.57427	.00564	-.33940	5.40*	†	-.03494	-.10939
Need for Help	.76026	.57800	.00373	.30565	3.57		.03274	.09359
% Receive Help	.76028	.57802	.00002	.04262	.02		-.00259	-.00607
Health Barriers	.76046	.57830	.00028	.03266	SNS		.00645	.04090
Reading Problems	.76073	.57871	.00041	-.05361	SNS		.00162	.01538
Behavior Barriers	.76376	.58333	.00462	-.16925	SNS		-.00995	-.06843
Math Specialist	.76383	.58344	.00011	.11540	.11	†	-.01047	-.03240
Team Teaching	.77379	.59875	.00256	-.14167	2.45		-.01446	-.04759
Open Space	.77214	.59620	.01276	-.22511	12.21***		-.02807	-.08745
Variability of Material Level	.77468	.60012	.00137	-.16028	1.31		-.00940	-.05237
Average Material Level	.78047	.60913	.00901	-.15091	8.62**	†	-.01003	-.19103

Note: SNS means that the individual F was not calculated because the F for the set was nonsignificant.

*p < .05

**p < .01

***p < .001

† Significant in 1975-76

Table V

Posttest Score Regression-All Grades: Contribution of Sets

Set No.	Set Name	Number of Variables in set	F		R^2 Change
1	Reading Pretest	1	436.45***	†	.456
2	Grade	1	72.97***	†	.076
3	Number of Test Levels	1	2.97	†	.003
4	Classroom Characteristics	3	0.70		.002
5	Teacher Characteristics	2	2.00		.004
6	Percent Receiving Free Lunch	1	5.05*		.005
7	Reading Behavior	4	7.38***	†	.031
8	Learning Barriers	3	1.69		.005
9	Classroom Environment	5	4.94***	†	.026

*p < .05
**p < .01
***p < .001
† Significant in 1975-76

the larger the percentage of students in a given classroom receiving free lunch, the lower are the posttest scores likely to be. It is a rough measure of the socioeconomic level of the class. The significance of the free-lunch variable suggests that classes with lower socioeconomic students tend to turn in lower performances on the posttest. It is important to note that this result is seen even after pretest effects are partialled out. Thus, classrooms in this sample with higher percentages of free-lunch children represented a somewhat disadvantaged learning environment, while those with lower percentages of free-lunch children were, by contrast, better learning situations.

The fourth significant set was the Reading Behavior group. Two individual variables within this set were significant. The first, Average Amount of Outside Reading, $F = 20.55$, $p < .001$, was based on the fact that students with higher posttest reading scores were observed by teachers to do more reading on their own. The second significant variable was average interest in reading, $F = 5.40$, $p < .05$. Results showed that students with high posttest scores also have higher interest in reading.

The final significant set, Classroom Environment, reflected the influence of two major variables. The first was the level of the most frequently used reading material in the classroom. The significant contribution, $F = 8.62$, $p < .01$, suggests that the lower the level of the most frequently used reading material, the greater the posttest score. The second significant variable was that dealing with the use of open space versus self-contained classrooms, $F = 12.21$, $p < .001$. The direction of the result indicates that self-contained classrooms are associated with higher posttest scores than are open space classrooms.

b. Discussion and Comparison. The regression on posttest scores across all grades showed that five major sets of variables are importantly involved in the prediction of these criterion scores. The first two, Reading Pretest

Score and Grade, are control variables. The significance of the grade variable is based on the fact that there was a decrease in posttest scores with increasing grade level. This result is consistent with that seen in the 1975-76 data and probably does not reflect an actual decrease in reading ability across grades, but rather a change in the characteristics of the test across its increasing levels. A third significant set of variables entering the equation, those dealing with Reading Behaviors, was also consistent with the 1975-76 result. Present data, as well as the earlier data, indicate that classrooms in which children do more reading on their own and are seen as having higher interest in reading have higher posttest scores. The consistency of this result across the two-year period is, itself, significant, since it indicates a probable stability of relationship among these variables.

A fourth significant set, Classroom Environment, suggests that certain aspects of the learning situation in which the child finds himself or herself are important. Here, the use of lower levels of reading material and self-contained, as opposed to open space, classrooms create a superior learning environment, at least to the extent that superiority is reflected in the reading posttest scores. While the classroom environment set was significant in the 1975-76 data, the nature of the significant variables within that set is not entirely consistent. Specifically, presence of a mathematics specialist was associated with higher scores in the earlier study, while here that variable was not significant, but the superiority of the self-contained classroom was shown. This difference in results probably reflects in large part a difference in the sample of classrooms available for analysis in the present study as compared with the earlier study. The larger, and hopefully more representative, sample of classrooms in the present study would suggest that current results are more likely to be stable and reliable than those from the 1975-76 data base. However, any result which is not consistent across the

two years must be viewed with caution, and it is recommended that specific administrative actions based on these inconsistent results be delayed until data for 1977-78 and later years can be examined and the likelihood of replication determined.

In addition to the inconsistencies noted above, two other results were inconsistent across the two-year period. The percentage of children receiving free lunches - our rough socioeconomic indicator - was not significant in the earlier data set, although we had expected that it would be. Here it was significant, indicating that classrooms with larger percentages of children receiving free lunches had lower posttest scores. In addition, the 1975-76 data suggested that more heterogeneous classrooms, in terms of the number of test levels found in the classroom, had higher average posttest scores. This heterogeneity (or homogeneity) variable was not significant in the 1976-77 data, and it will require further examination in future years before firm conclusions can be reached.

3. Regression Analyses for Primary Grades (Grades 1-3)

Data for the first three grades were entered into a separate regression analysis similar to the one described above.

a. Results. Means and standard deviations for the predictor and criterion variables are provided in Table I (Chapter IV). Tables VI and VII provide the regression summary for separate variables and the F-ratios and associated R^2-change values for sets, respectively, for the regression analyses of the primary grades. Table B (in Appendix E) provides the correlation matrix. The final, significant multiple correlation for the primary grades was .77, accounting for approximately 59% of the variance. The estimated population R^2 based on the latrer is..57.

Three sets of variables entered significantly into the final regression equation. They were, in order of entry into the hierarchical analyses, Pretest

Table VI

Posttest Regression Summary: Primary Grades

Variable	Multiple R	R Square	RSQ Change	Simple R	F-Ratio		B	Beta
Pretest Score	.68425	.46820	.46820	.68425	194.35***	†	.63709	.54008
Grade	.69326	.48061	.01241	.29420	5.15*		.00709	.04819
Test Level Count	.69740	.48636	.00575	-.15043	SNS		-.00525	-.03699
Class Size	.70290	.49407	.00771	.11185	SNS		.00267	.08815
% Males	.70292	.49409	.00003	-.11864	SNS		.00082	.05899
Class Holding Rate	.70335	.49470	.00061	.01387	SNS		-.00036	-.02196
Teacher Turnover	.70776	.50092	.00622	-.04826	SNS		-.07217	-.08678
Teacher Absence	.70911	.50284	.00192	-.06422	SNS		-.18486	-.05269
% Free Lunch	.71337	.50890	.00606	-.34838	SNS		-.02123	-.05436
Reading on Own	.74827	.55991	.05101	-.44590	21.17***	†	-.05286	-.10343
Reading Interest	.75253	.56630	.00639	-.40234	2.65	†	-.03867	-.11980
Need for Help	.75446	.56921	.00291	.40834	1.21		.02007	.05672
% Receive Help	.75459	.56941	.00020	.07635	.08		-.00438	-.00959
Health Barriers	.75470	.56958	.00017	-.01341	SNS		.00470	.03086
Reading Problems	.75472	.56960	.00002	-.05432	SNS		.00184	.01755
Behavior Barriers	.76159	.58002	.01042	-.18132	SNS		-.01561	-.09804
Math Specialist	.76347	.58288	.00287	.21541	SNS		.01445	.03854
Team Teaching	.76892	.59124	.00003	-.27983	SNS		.00564	.01861
Open Space	.76890	.59121	.00833	-.34416	SNS		-.03521	-.11265
Variability of Material Level	.76959	.59226	.00102	-.17783	SNS		-.00659	-.03729
Average Material Level	.76997	.59285	.00059	.20486	SNS		-.00257	-.03492

Note: SNS means that the individual F was not calculated because the F for the set was nonsignificant.
*p < .05
**p < .01
***p < .001
† Significant in 1975-76

Table VII

Posttest Score Regression - Primary Grades: Contribution of Sets

Set No.	Set Name	Number of Variables in Set	F		R^2 Change
1	Reading Pretest	1	194.35***	†	.468
2	Grade	1	5.15*		.012
3	Number of Test Levels	1	2.39		.006
4	Classroom Characteristics	3	1.15		.008
5	Teacher Characteristics	2	1.70		.008
6	Percent Receiving Free Lunch	1	2.52		.006
7	Reading Behavior	4	6.28**	†	.061
8	Learning Barriers	3	1.47		.281
9	Classroom Environment	5	1.07		.013

*p < .05
**p < .01
***p < .001
† Significant in 1975-76

Score, Grade, and Reading Behavior. The Pretest set made a contribution similar to that seen in the All Grades analyses above. As would be expected, classrooms with higher average pretest scores had higher average posttest scores. Again, however, this variable was entered in as a control variable in order to partial out its affects before examining other variables. The second significant set, was Grade. Here again posttest scores decreased as grade level increased, reflecting differences in the test itself. The third and final significant set, Reading Behaviors, contained one significant variable. This variable, Class Average Amount of Outside Reading, $F = 21.17$, $p < .001$, reflected higher posttest scores in classes with larger amounts of outside reading. The remaining variables in the set, including reading interest, need for help and percentage of students receiving help, were not significant.

b. Discussion and Comparison. The regression analyses for primary grades showed, overall, that two sets of variables are predictive of reading achievement at this level. The first of these, pretest score, was a control variable and needs no further discussion. The second, the reading behavior set, was also found to be significant in the 1975-76 analyses reported earlier. However, in these earlier analyses, the amount of interest in reading found in students in a given classroom, as well as the average amount of outside reading, was predictive of achievement scores. The partial inconsistency in the reading behavior data between the two years studied suggests caution in interpreting the reading interest variable. Examination of data for 1977-78 and later years will be essential in reaching conclusions concerning this variable.

4. Intermediate-Grade Regression Analyses (Grades 4-6)

The final regression analysis involving posttest achievement scores was for the three intermediate grades. The predictor variables entered were the

same as those in the two analyses reported above.

a. Results. Means and standard deviations for the intermediate grade variables are provided in Table I (Chapter IV). Tables VIII and IX provide, respectively, the regression summary of the F-ratios for sets and statistics for individual variables. Table C (in Appendix F) provides the correlation matrix. Results of this analysis showed significant contributions of only two predictor sets. These were the pretest control variable and grade level. Pretest level performed as expected and as in previous analyses, higher pretest scores predicting higher posttest scores. Grade level was consistent with the all-grades and primary Grades analyses in showing a decrease in posttest scores with increasing grade levels. No other sets of variables in the intermediate grade analyses were significant.

b. Discussion and Comqparison. Results of the regression analysis of the posttest scores in the intermediate grades were straightforward. They showed that, in the 1976-77 year, only the reading pretest and grade control variables significantly predicted posttest reading scores. None of the other variables entered into the study were able to account for substantial amounts of variance. These results are not entirely consistent with those for the 1975-76 year. There, two additional sets, number of test levels present in the classroom and classroom environment, were significant. The first of these defines the degree of heterogeneity present in a given class, while the second examines the effects of a number of variables relevant to the nature of the learning situation surrounding the student. It might be noted that this latter set did closely approach significance in the 1976-77 data for intermediate grades and was, of course, significant in the All Grades analysis.

5. Overview of Posttest Score Results

Overall results for the posttest score regressions suggest, as in the 1975-76 data, the importance of controlling for the effect of pretest score

Table VIII

Posttest Regression Summary: Intermediate Grades

Variable	Multiple R	R Square	RSQ Change	Simple R	F-Ratio		B	Beta
Pretest Score	75438	.56909	.56909	.75438	289.39***	†	.68617	.69012
Grade	.76179	.58032	.01123	-.08094	5.71*	†	-.00582	-.04198
Test Level Count	.76664	.58773	.00742	.18510	SNS	†	.01534	.13268
Class Size	.76676	.58792	.00019	-.00973	SNS		-.00145	-.06746
% Males	.76728	.58871	.00079	-.04393	SNS		-.00060	-.05611
Class Holding Rate	.76904	.59142	.00271	.17668	SNS		.00013	.00879
Teacher Turnover	.76932	.59185	.00043	.02408	SNS		-.03723	-.05266
Teacher Absence	.77118	.59471	.00287	-.02275	SNS		-.09910	-.05009
% Free Lunch	.77542	.60128	.00656	-.26802	SNS		-.03450	-.09566
Reading on Own	.77717	.60399	.00272	-.16546	SNS		.00622	.01446
Reading Interest	.77954	.60768	.00369	-.24059	SNS		-.01210	-.04301
Need for Help	.78499	.61621	.00853	.27352	SNS		.03802	.12414
% Receive Help	.78586	.61758	.00137	.10870	SNS		.00917	.02554
Health Barriers	.78663	.61879	.00121	-.01007	SNS		-.00129	-.00872
Reading Problems	.78681	.61907	.00028	-.13318	SNS		-.00164	-.01741
Behavior Barriers	.78704	.61943	.00036	-.13405	SNS		-.00649	-.05467
Math Specialist	.78759	.62029	.00086	.14343	SNS	†	-.01928	-.07499
Team Teaching	.79049	.62487	.00346	-.07325	SNS		-.02026	-.07443
Open Space	.78830	.62142	.00112	-.20890	SNS		.00073	.00244
Variability of Material Level	.79225	.62766	.00278	-.22626	SNS		-.02063	-.12700
Average Material Level	.80008	.64013	.01247	.08116	SNS	†	-.01569	-.19684

Note: SNS means that the individual F was not calculated because the F for the set was nonsignificant.

*p < .05
**p < .01
***p < .001
† Significant in 1975-76

Table IX

Posttest Score Regression - Intermediate Grades: Contribution of Sets

Set No.	Set Name	Number of Variables in Set	F		R^2 Change
1	Reading Pretest	1	289.39***	†	.569
2	Grade	1	5.71*	†	.011
3	Number of Test Levels	1	3.77	†	.007
4	Classroom Characteristics	3	0.625		.004
5	Teacher Characteristics	2	0.837		.003
6	Percent Receiving Free Lunch	1	3.34		.007
7	Reading Behavior	4	2.07		.016
8	Learning Barriers	3	0.31		.002
9	Classroom Environment	5	2.11	†	.031

*p < .05
**p < .01
***p < .001
† Significant in 1975-76

and grade level. The variable representing the number of test levels present in the classroom, and hence the heterogeneity of grouping of students, was not significant in any of the achievement analyses in the present study. This interesting variable was significant in the all-grades analyses and in the analyses for intermediate grades in the 1975-76 year, where results suggested that more heterogeneous classrooms have higher achievement scores. It was suggested in the earlier report that results for this variable should be interpreted with some caution, and present results suggest that the caution was justified. The larger sample of classrooms in the present study did not confirm the significance of the test level variable. Since results for the two years are inconsistent, it will be necessary to further examine the data from both these years and also to await further data from 1977-78 and later years before drawing firm conclusions as to the likely role of reading level heterogeneity in the classroom.

The set of Reading Behavior variables was significant in both the All-Grades and Primary Grades analyses in the present study, suggesting, in particular, the importance of amount of outside reading and average interest in reading in relation to posttest scores. It is of particular interest that these results are almost entirely consistent with those in the 1975-76 data for the reading behavior variable set. Unfortunately, it is not yet possible to determine the exact nature of the relationship implied by these results. In particular, cause-effect relationships cannot be directly determined from available data. On the basis of present results, however, it would certainly not be inappropriate to take such steps as may be possible to increase the students' level of interest in reading and encourage them to do increased amounts of outside reading. Making available to the children interesting school materials for home use, active encouragement of reading, and cooperation with area libraries in encouraging outside reading programs for children,

would be examples of the kinds of activities that one might hope, on the basis of present results, would increase posttest reading scores. Again, however, it must be cautioned that cause-effect relationships cannot be entirely determined from present data.

The results show some consistency across years in the role of certain classroom environment variables in reading performance. Of particular interest is the fact that the level of reading materials employed in the classroom appears to be related to posttest reading performance. The result is somewhat complicated by the fact that the classroom environment set did not quite reach significance in the 1976-77 data set. However results for 1975-76 and for the all-grades analyses in the 1976-77 combined with the nearly significant result for the intermediate grades would suggest the need to attend closely in further work to the level of materials used in the classroom. What results suggested is that, at the intermediate grade level, classrooms which use lower levels of reading materials tend to achieve higher posttest scores. This may suggest that where the teacher "overshoots" the reading ability of students by using materials too complex for them to understand, the long-term result is detrimental, at least in terms of year-end test scores. The other variables in the classroom environment set were not consistent across years. The math specialist variable was significant in 1975-76, while the open-space versus self-contained class variable was significant in the present data but not in the earlier year. The latter variable is, however, of particular interest, and it would seem to be of value to further analyze the impact of this variable within present data and to attend closely to its possible significance in later years.

C. Regression Analysis of Reading Interest

As has been noted, a series of three hierarchical regression analyses, paralleling those done on posttest performance, were undertaken with regard

to teacher ratings of interest in reading. Thus, for these analyses, the criterion variable was a reading interest score based on teacher ratings as described above. The All-Grades, primary grades and intermediate grades analyses are reported below.

1. Basic statistics.

Table I (in Chapter IV) provides the complete set of means and standard deviations for the criterion and predictor variables.

2. Analysis of All Grades.

Data from all six grades were entered into the overall regression analyses involving the total of 396 classes across the 34 pilot schools. A presentation of the statistical results and a discussion of these results follows.

a. Results. Table A (Appendix D) provides the correlations. Table X provides a summary of the results of the all-grades regression. Table XI summarizes the F-ratios and the R^2 change values for the nine sets of predictor variables employed. It should again be noted that these sets are not identical to those employed in the posttest regressions. Specifically, Set 4, Classroom Characteristics, includes class absence rate, a variable not used in the posttest regression. In addition, the reading interest variable, which was a predictor in posttest analyses, is, of course, now the criterion variable. Finally, Average Amount of Reading, was dropped from the present analysis, since it is highly correlated with Reading Interest and was rated by teachers at the same time they were rating Reading Interest, leading us to suspect a halo effect. This would make the Amount of Reading variable one of questionable utility as a predictor of reading interest.

The final multiple correlation for the all-grades regression was .52, accounting for approximately 27% of the variance. Based on this latter value, the more conservative estimate of the population R^2 is .26.

Table XI shows that five of the nine sets of variables contribute significantly to the final regression equation. In order of entry into the

Table X

Reading Interest Regression Summary: All Grades

Variable	Multiple R	R Square	RSQ Change	Simple R	F-Ratios		B	Beta
Pretest Score	.23531	.05537	.05537	-.23521	28.99***	†	-.00360	-.09566
Grade	.34118	.05817	.00280	.03483	SNS		.03119	.13407
Test Level Count	.25251	.06376	.00559	.07189	SNS		-.00290	-.00702
Class Size	.25256	.06378	.00002	.01274	.01	†	-.00000	-.00005
% Males	.29091	.08463	.02085	.17495	10.92***	†	.00615	.06452
Class Holding Rate	.32346	.10463	.02000	.17841	10.47**		.00545	.13441
Teacher Turnover	.34378	.11819	.01356	-.10784	7.10**	†	-.00277	-.05409
% Free Lunch	.34599	.11971	.00152	-.03843	SNS	†	-.08089	-.03165
Reading on Own	.35044	.12281	.00310	.04247	SNS		.21339	.02537
Reading Interest	.35861	.12860	.00579	.17606	SNS		.06115	.04895
Need for Help	.45674	.20861	.08001	-.39659	41.89***	†	-.34924	-.31890
% Receive Help	.46916	.22011	.01150	.04450	6.02*	†	.12128	.09081
Health Barriers	.47502	.22565	.00554	-.02211	2.90		-.03837	-.07772
Reading Problems	.48369	.23395	.00831	.00334	4.35*		-.04456	-.13519
Behavior Barriers	.48893	.23905	.00510	.14375	2.67		.03754	.08249
Math Specialist	.49417	.24420	.00515	-.14267	2.70		-.04907	-.04850
Team Teaching	.50926	.25935	.01502	.15539	7.86**		.13763	.14471
Open Space	.49430	.24433	.00013	.10890	.07		-.10441	-.10387
Variability of Material Level	.51074	.26086	.00151	-.06456	.79		-.00940	-.05719
Average Material Level	.52361	.27417	.01331	.22763	6.97**		.07423	.13207

Note: SNS means that the individual F was not calculated because the F for the set was nonsignificant.

*p < .05
**p < .01
***p < .001
† Significant in 1975-76

Table XI

Reading Interest Score Regression - All Grades: Contribution of Sets

Set. No.	Set Name	Number of Variables in Sets	F		R^2 Change
1	Reading Pretest	1	28.99***	†	.554
2	Grade	1	1.47		.003
3	Number of Test Levels	1	2.93		.006
4	Classroom Characteristics	4	7.12***	†	.054
5	Teacher Characteristics	2	1.21	†	.005
6	Percent Receiving Free Lunch	1	3.03	†	.006
7	Reading Behavior	2	23.96***	†	.092
8	Learning Barriers	3	3.31*		.019
9	Classroom Behavior	5	3.68**		.035

*p < .05
**p < .01
***p < .001
† Significant in 1975-76

hierarchical analysis, the first set to contribute significantly was the reading performance pretest score. The second set, classroom characteristics, contained three individual variables which contributed significantly to the prediction of interest ratings. These were Class Absence rate, $F = 10.92$, $p < .001$, percentage of male students in the class, $F = 10.47$, $p < .001$, and class holding rate, $F = 7.10$, $p < .01$. Directional relationships indicated that interest scores are higher in classes with low absence rates, lower percentages of male students, and high holding rates.

Reading Behavior was the third set. Both the perceived need for extra help in reading, $F = 41.89$, $p < .001$, and the percentage of students receiving help, $F = 6.02$, $p < .05$, were significant. Results showed that classrooms having a high percentage of children who need help as well as those who have a high percentage of children receiving help tend to have lower reading interest scores. The fourth set, Learning Barriers, showed a reliable effect of reading problems, $F = 4.35$, $p < .05$. Classes with larger numbers of children with reading problems have lower reading interest. Finally, the classroom environment set was significant. Two individual variables within that set, the presence or absence of team teaching, $F = 7.86$, $p < .01$, and the level of most often used reading material, $F = 6.97$, $p < .01$, were significant. The team teaching variable shows that classes are more likely to have high interest in reading if they are involved in a team-teaching situation. The second variable, level of most often used material, shows that classes using lower levels of materials are likely to have higher interest in reading.

b. Discussion and Comparison. The reading interest and regression analyses showed, first, that it is possible to predict reading interest from the variables available for study in the present analyses. Aside from the control variable, reading pretest score, the major contributors are classroom

characteristics, reading behaviors, learning barriers, and classroom environment.

The classroom characteristic set contained three reliable predictors of reading interest. Two of these, lower absence rate and higher class holding rate, are not particularly surprising and would seem to suggest that encouraging the stability of class membership over the course of the year and reducing absence rates would be desirable. The third variable suggests that classes with lower percentages of males may be expected to exhibit higher reading interest levels. This particular variable did not contribute significantly in the 1975-76 study and should thus be treated with caution until confirmed by data from later years. If confirmed, it may reflect a sociocultural phenomenon in which males are encouraged to be less and females more interested in early academic pursuits. The fourth variable in the classroom characteristic set, class size, was significant in the 1975-76 data, results there suggesting higher reading interest in larger classes. The fact that the variable did not hold up in the second study suggests considerable caution in interpreting its significance. Results for the reading behavior set were entirely consistent with those in the 1975-76 study. Both variables in the set were, both in the present and the earlier study, reliable predictors of reading interest. This suggests a need for interpretation of and close attention to these reading help variables, both in future studies and in consideration of school policy. Some alternative interpretations of the possible meanings of the significance of these variables are discussed in reporting the analysis of primary grade data below.

The final two significant sets were those dealing with learning barriers and classroom environment. Each set contained a single significant variable, reading barriers in the first case and level of most used reading material in the second. It is perhaps not surprising to learn that classes with large

numbers of students referred for reading problems have lower interst in reading. However, it is notable that this variable was not a significant predictor in the earlier study.

The significance of the team teaching variable in the environment set is of interest. What this variable suggests, essentially, is that team teaching enhances, or at least is positively associated with, increased interest in reading. Since it has already been seen, both in the present study and in the earlier study, that reading interest may be associated with increased posttest scores, this finding may have important policy implications. Unfortunately, it is probably too early to draw firm conclusions to this effect, since the team teaching variable was not significant in the 1975-76 study. The variable should be examined further in future years before it is used as a firm basis for actual policy changes.

Two sets of variables found to be significant in the earlier study were not significant for the all-grades interest analysis in the present study. These were teacher characteristics, within which teacher turnover rate was a significant variable in the earlier study, and percentage of students receiving free lunch. Since neither of these variables held up in the present study, they will require further examination in data for future years before firm conclusions can be reached.

3. The Analysis of Primary Grades

The means and standard deviations for predictor and criterion variables are provided in Table I above (Chapter IV).

a. Results. Tables XII and XIII provide the regression summary for variables and sets in the analysis of primary grade data, and Table B (Appendix E) provides the correlation matrix.

The final multiple correlation of .58 accounted for approximately 33% of the variance. A more conservative estimate of the population R^2 suggests a

Table XII

Reading Interest Regression Summary: Primary Grades

Variable	Multiple R	R Square	RSQ Change	Simple R	F-Ratios		B	Beta
Pretest Score	.27187	.07391	.07391	-.27187	18.85***	†	-.00428	-.11178
Grade	.29672	.08804	.01413	-.18757	SNS		-.03330	-.07300
Test Level Count	.30167	.09101	.00297	-.00174	SNS	†	-.06073	-.13801
Class Size	.30554	.09335	.00234	.02987	.60	†	-.00107	-.01138
% Males	.30650	.09394	.00059	.11790	.15		-.00215	-.02142
Class Holding Rate	.36289	.13169	.03774	.24163	9.63**		.00732	.17005
Teacher Turnover	.39134	.15315	.02146	-.11337	5.47*		-.00395	-.07781
% Free Lunch	.39176	.15347	.00033	-.00010	SNS		.08675	.03366
Reading on Own	.39192	.15360	.00013	.00288	SNS		-.11537	-.01061
Reading Interest	.39193	.15361	.00001	.11057	SNS		-.01371	-.01132
Need for Help	.47151	.22233	.06871	-.39492	17.53***	†	-.37006	-.33748
% Receive Help	.50102	.25102	.02869	.07417	7.32**		.19778	.13989
Health Barriers	.51475	.26497	.01395	-.01954	SNS		-.04508	-.09553
Reading Problems	.52405	.27463	.00966	-.05626	SNS		-.06081	-.18746
Behavior Barriers	.52912	.27997	.00534	.13382	SNS		.04441	.09004
Math Specialist	.54588	.29799	.01802	-.21227	4.60*		-.16924	-.14567
Team Teaching	.56591	.32025	.02110	-.20912	5.38*		-.16854	.17954
Open Space	.56494	.29915	.00116	.13936	.30		-.17329	.17896
Variability of Material Level	.56591	.32025	.00000	-.18086	0.00		-.00219	-.00961
Average Material Level	.57754	.33356	.01331	.16418	3.40		.07746	.14140

Note: SNS means that the individual F was not calculated because the F for the set was nonsignificant.

*p < .05
**p < .01
***p < .001
† Significant in 1975-76

Table XIII

Reading Interest Score Regression - Primary Grades: Contribution of Sets

Set No.	Set Name	Number of Variables in Set	F		R^2 Change
1	Reading Pretest	1	18.85***	†	.074
2	Grade	1	3.60		.014
3	Number of Test Levels	1	0.76	†	.003
4	Classroom Characteristics	4	3.96**	†	.062
5	Teacher Characteristics	2	0.06		.0005
6	Percent Receiving Free Lunch	1	0.003		.0001
7	Reading Behavior	2	12.42***	†	.097
8	Learning Barriers	3	2.46		.029
9	Classroom Environment	5	2.73*		.054

*p < .05
**p < .01
***p < .001
† Significant in 1975-76

value of approximately .30.

There were four sets of variables that significantly contributed to the prediction of the reading interest score. The first was the control variable, pretest score. The second was the classroom characteristic set. Two variables within this latter set, percentage of males in the class and class holding rate, were individually significant. The significance of the sex distribution variable, $F = 9.63$, $p < .01$, was due to the fact that classes with higher percentages of males tended to have more interest in reading, the same result found in the all grades analyses. The class holding rate variable, $F = 5.47$, $p < .05$, showed that classes with a high turnover of students tended to have significantly lower levels of interest in reading.

The remaining two significant sets were reading behavior and class environment. The reading behavior set contained two significant variables, the average need for help in reading, $F = 17.53$, $p < .001$, and the average percentage of students receiving help, $F = 7.32$, $p < .001$. Basically, these two variables indicate that the greater the perceived need for extra help in reading and the higher the percentage of students receiving help elsewhere, the lower the reading interest level in the room. The two significant variables in the classroom environment set were the presence or absence of a math specialist, $F = 4.60$, $p < .05$, and the presence or absence of a team teaching situation, $F = 5.38$, $p < .05$. Basically, high interest classes were those in which a math specialist was available and team teaching was employed.

b. Discussion and Comparison. Results of the regression analysis concerned with factors contributing to reading interest level in the primary grades showed that certain class characteristics, as well as reading behavior and classroom environment, may be important. The class characteristic set, significant in the present study, was also significant in the study of 1975-76 classes. However, in the earlier year the major variable within this set

contributing to the prediction of interest scores was class size, while the major predictors in the present study were sex distirbution and class holding rate. Results across the two years and across the all-grades and primary grades analyses thus indicate that some subset of class characteristics is of importance in predicting reading interest scores. However, the specific nature of the most important class characteristics is unclear. In the present study, both sex distribution and class holding rate are important variables at the primary grade level. This would suggest the need to attend to the sex distribution factor in attempts to increase reading interest levels in the primary grades. The class holding rate variable appears to suggest the need to assure, as much as possible, the stability of class membership over the course of the school year, although the exact nature of cause-effect relationships cannot be entirely determined from present data. As with some other results reported above, firm conclusions concerning the role of class characteristics in predicting reading interest will have to be held in abeyance until further studies can be done on data for future years, since the specific class characteristics variables contributing significantly to the regression equation were different in the present study than in the previous study.

The reading behavior set, significant in the present study, was also significant in the 1975-76 study. In particular, one of the two variables within that set, the perceived need for help in reading, was individually significant across the two-year period. This may reflect a relationship between reading skills and reading interest, in that classes with lower levels of reading skills (and hence needing more help with reading) may lack interest in reading. The relationship has often been assumed to be a reciprocal and interacting one. That is, children are not interested in doing something that they do not do well (in this case reading). However, the child's lack

of skill in reading may reflect an earlier lack of interest in the subject. In interpreting the possible effects here, it is also important to keep in mind that both the reading interest variable and the need-for-help variable are based on the perceptions of the teacher. An alternative interpretation of this result, then, may be that teachers who perceive their classes as needing high amounts of extra help in reading also tend to perceive those classes as lacking interest in reading. Whatever the interpretation, the consistent significance of the reading help variables in relationship to reading interest would suggest the need to attend to these variables closely in the future.

The two significant classroom environment variables, presence of a math specialist and team teaching, must be considered a preliminary indication of the possible relevance of these variables to reading interest. The presence of a math specialist related, in the 1975-76 data to reading posttest achievement performance, particularly in the intermediate grades, but did not relate significantly to interest at any level. The team teaching variable was not related, in the earlier study, to either achievement or interest. Thus, these variables should definitely be studied further before firm conclusions are drawn.

4. Analysis of Intermediate Grades

A. The analysis for the three intermediate grades involved a total 121 classes with the same predictor variables as in the preceding two analyses being entered into the hierarchical regression analysis.

a. Results. Relevant means and standard deviations for the predictor and criterion variables appear in Table I (Chapter IV). Table C (Appendix F) provide the correlation matrix. Tables XIV and XV contain, respectively, the regression summary for individual variables and the regression summary for sets.

Table XIV

Reading Interest Regression Summary: Intermediate Grades

Variable	Multiple R	R Square	RSQ Change	Simple R	F-Ratios		B	Beta
Pretest Score	.20077	.04031	.04031	-.20077	10.31**	†	-.00322	-.08776
Grade	.20128	.04051	.00021	.00794	SNS		-.00897	-.01821
Test Level Count	.24145	.05830	.01778	.10041	4.55*	†	.02776	.06757
Class Size	.25004	.06252	.00422	-.02434	1.08		-.00096	-.01255
% Males	.32872	.10805	.04554	.24084	11.64***	†	.00981	.10878
Class Holding Rate	.34041	.11588	.00782	.12031	2.00	†	.00484	.12728
Teacher Turnover	.35777	.12800	.01212	-.13002	3.10		-.00163	-.03129
% Free Lunch	.36813	.13552	.00752	-.07681	SNS	†	-.17105	-.06809
Reading on Own	.37370	.13965	.00413	.06103	SNS		.22351	.03179
Reading Interest	.39651	.15722	.01757	.25193	4.49*	†	.12913	.10075
Need for Help	.48546	.23567	.07845	.41586	20.04***	†	-.32987	-.30305
% Receive Help	.48624	.23643	.00075	-.00665	.18	†	.02588	.02028
Health Barriers	.48624	.23643	.00001	.08614	SNS		-.01031	-.01961
Reading Problems	.49304	.24309	.00666	.08648	SNS		-.03380	-.10080
Behavior Barriers	.49378	.24382	.00073	.14333	SNS		.01711	.04058
Math Specialist	.49623	.24624	.00242	-.11719	SNS		-.00379	-.00414
Team Teaching	.50785	.25791	.01163	.12139	SNS		.11339	.11724
Open Space	.49626	.24627	.00003	.10393	SNS		-.05401	-.05102
Variability of Material Level	.50945	.25954	.00163	-.24656	SNS		.01250	.04412
Average Material Level	.52898	.27982	.02029	.31835	SNS		.10386	.17994

Note: SNS means that the individual F was not calculated because the F for the set was nonsignificant.

*p < .05
**p < .01
***p < .001
† Significant in 1975-76

Table XV

Reading Interest Score Regression - Intermediate Grades: Contribution of Sets

Set No.	Set Name	Number of Variables in Sets	F		R^2 Change
1	Reading Pretest	1	10.30**	†	.040
2	Grade	1	.05		.0002
3	Number of Test Levels	1	4.55*	†	.018
4	Classroom Characteristics	4	4.45**	†	.070
5	Teacher Characteristics	2	1.49	†	.012
6	Percent Receiving Free Lunch	1	4.49*	†	.018
7	Reading Behavior	2	10.12***	†	.080
8	Learning Barriers	3	.629		.007
9	Classroom Environment	5	1.84		.036

*p < .05
**p < .01
***p < .001
† Significant in 1975-76

The final multiple correlation of .53 accounted for approximately 28% of the variance. The more conservative R^2 estimate for the population is .25.

For the intermediate-grade analyses, five sets of variables entered significantly into the regression equation. As can be seen in Table XV, they were reading pretest score, number of test levels seen in the classroom, classroom characteristics, percentage of students receiving free lunch, and reading behaviors. The pretest score, as a control variable, requires no comment. The test level count variable indicates that greater heterogeneity in the classroom (larger numbers of test levels present) is associated with lower reading interest. The third set, classroom characteristics, contains one variable, class absence rate, that is significant, $F = 10.64$, $p < .001$. The higher the class absence rate, the lower the interest in reading. The third set, percentage of children receiving free lunch, showed that the higher the percentage of children receiving free lunch, the lower the level of interest in reading in a given classroom. Finally, the reading behavior set was reliably associated with reading interest. Only one variable in the set was individually significant. This variable, the perceived need for help, $F = 20.04$, $p < .001$, indicates, as in the all-grades and primary grades analyses, that classrooms perceived as needing larger amounts of extra help in reading are also perceived as having lower interest in reading.

b. Discussion and Comparison. The test level count variable, significant both in the present study and in that involving the 1975-76 data, suggests that homogeneous grouping is associated with higher interest in the intermediate grades. This finding might have important policy implications, particularly when combined with the observation that greater homogeneity of grouping was associated with higher posttest scores in the 1975-76 data. The fact that this latter finding was not replicated in the present study must make our conclusions concerning the test level count variable somewhat more

tenuous. However, it is a variable that certainly bears further attention and consideration.

The significant class absence rate variable in the classroom characteristics set was also consistent with the intermediate grades analysis for 1975-76. Higher class absence rates are associated with lower interest in reading, quite possibly suggesting that increasing the level of interest for students in the classroom will affect both reading and class absences. The significance of the socioeconomic indicator, percentage of students receiving free lunch, is also consistent with the finding for the earlier year. Again, classes with higher percentages of children receiving free lunches tend to have lower interest in reading. This is not a surprising finding, and it is difficult to recommend any particular policy changes that might alleviate the low interest problem in classes with low socioeconomic children.

The final significant set, reading behavior, showed that, once again, the classes perceived as having a high need for help in reading were also perceived as having low interest in reading. This result was consistent with the 1975-76 study and has been discussed in detail above in connection with the analysis of data for the primary grades.

VI. EVALUATIONS OF READING PROGRAMS

A. Methods

In this chapter evaluations of reading programs (as defined in Chapter IV D) are reported. The methods used were analysis of variance (ANOVA) and analysis of covariance (ANCOVA). Based on a set of selection criteria, types of programs were compared using the Posttest % Objectives Mastered on the Prescriptive Reading Test (PRT) as the dependent variable. These analyses, unlike the regression analyses discussed in the preceding chapter (Chapter V) used the student as the unit of analysis. That is, data for all students having the same type of reading program (as defined in Chapter IV D) were grouped together regardless of whether or not students came from the same classes or even the same schools. For both sets of analyses, students were drawn from all classes at a given grade level, with the lower and upper dividing points being .5 or more and .4 or less to include multi-grade classes. That is, for example, all students in classes whose computed unweighted average class level was between 3.5 and 4.4 were classified as 4th grade students, regardless of their true grade level.

Two different analyses were done for each grade level. The first was a one-way analysis of variance (ANOVA) and covariance (ANCOVA) in which the independent variable was reading programs. The second analysis was a two-way analysis of covariance (ANCOVA) in which the independent variables were reading programs and whether or not the student received extra help in reading. The dependent variable in both sets of analyses was the percentage of posttest reading objectives mastered. The one-way analysis was done with the BMDP1V ANCOVA computer program; the second used the BMDP2V ANCOVA program.

For both sets of ANCOVA (one-way, two-way), two covariates were used: pretest percentage of reading objectives mastered, and the coded reading level

of the first reading materials (i.e., of the reading material used most frequently by the student). For both sets of analyses at a given grade level, the same reading programs and samples of students were used.

The same analyses were also done using a logit transformation of the two test variables. If the pretest or posttest percentage is converted to a proportion, p, the logit is: $\log (p + 1/6N) - \log (1 - p + 1/6N)$ (Mosteller & Tukey, 1977). The N in the expression is the total number of objectives of the test. The descriptive data reported in the tables to follow are the percentage data. The significance tests reported are for the logit transformation analyses. The analyses of transformed pre and post measures gave results that were largely, but not entirely, similar to results for the untransformed variables. The logit transformation has the effect of spreading out the percentage scale increasingly the further the departure in either direction from 50% (or 0 logit), and of stabilizing variances.

The criteria for selecting types of reading programs to be compared for both sets of analyses were that the program had 50 or more students, and that losses (from missing variate or covariate data) left sample sizes of 30 or more. Programs (more correctly, types of programs) meeting those criteria at each grade level are identified by an asterisk in Table III (Chapter IV).

In the one-way analyses of covariance, if a test of the equality of regression slopes of the groups indicated heterogeneity (at the .05 level of significance), no further significance tests were interpreted. If slopes were homogeneous, but the test of the equality of adjusted group means was not significant at the 5% level of chance, no other statistical tests were interpreted. In the two-way ANCOVA, the test for interaction is the test of homogeneity of regression slopes.

The one-way set of analyses of covariance was made to test the null hypothesis that there was no difference in the effect of different reading programs on

achievement. The two-way set was made to test the null hypotheses that there were no differences in effects of different reading programs and in receiving extra help in reading. There are assumptions underlying the use of analysis of covariance with the data reported here. One is that the effects of treatments are the same regardless of the starting levels (pretest) of different groups. Starting levels of groups were generally not equal. There are other assumptions involved (conceptual and technical) in the analyses. We thus believe that results should be viewed with some caution. It will be found, however, that analyses of data from the prior year ('75-'76) will help to establish what conclusions can be drawn with greater and lesser confidence. In cases in which regression slopes were not equal (homogeneous), further analyses of the groups, examining regions of significant differences in slopes, were made to help clarify the differential effects of different programs.

The reader is reminded that different students in a particular reading program (as defined) as a given grade level may have taken different levels of the PRT. Using percentages of objectives mastered as the measure, rather than the raw number of objectives, was the means employed to try to equalize scales for different test levels as much as possible short of conventional norming.

B. Comparison of the First Grade Reading Programs

Six types of reading programs met inclusion criteria:

- Basal only (including Basal-Basal) materials
- Basal-Linguistics/Phonics materials
- Basal-Language Experience materials
- Basal-Programmed/Structured materials
- Linguistics/Phonics
- Linguistics/Phonics-Supplementary materials

The one-way analysis of variance test of unadjusted group means was significant ($F_{5,497} = 12.406$, $p = .0000$); the one-way analysis of covariance test of

the equality of adjusted group means was also significant ($F_{5,495} = 12.069$, $p = .0000$). Regression slopes were homogeneous ($F_{10,485} = 1.082$, $p = .3745$). Table XVI shows for each group the unadjusted means and standard deviations, the pre-post difference of the unadjusted means, the adjusted means, and regression slopes for each reading program, based on the analysis of untransformed variables. It also shows the rank order of the unadjusted and adjusted post-test means for each group relative to the rest of the groups (1 = highest, 6 = lowest), based on the analysis of logit transformations.

Table XVII gives similar data for the two-way ANCOVA of reading programs and receiving or not receiving extra help in reading. In this analysis there was a significant effect of reading programs ($F_{5,489} = 4.791$, $p = .000$), but not of receiving extra help ($F_{1,489} = .060$, $p = .806$). There was a significant interaction ($F_{1,489} = 3.045$, $p = .010$).

Table XVIII shows the number of classes and number of students per class per program going into these analyses. The two analyses suggest that the type of reading program used at the first grade level did seem to have different effects in general. The fact that regression slopes were equal in the one-way ANCOVA suggests that relationships for that analysis were directly interpretable.

In the main comparison of reading programs (the one-way ANCOVA) at the first grade level, the Basal-Language Experience type of program was significantly more effective on the average than the other types for the students in the sample analyzed. The other types of programs were not significantly different from each other, on the average, in effectiveness with students throughout the pretest range.

The two-way analysis suggests that receiving extra help in reading did not seem to make an appreciable difference in reading achievement, although since we have reason to suspect that relationships may be more complex, a more detailed analysis would be appropriate. Thus, we can conclude only that

Table XVI

Statistical Data for Each Reading Program Compared, Grade 1

Program	Variable	Students	Mean	SD	Adjusted Mean	Regression Slope+	Rank of:+ Unadj. M.	Rank of:+ Adj. M.
Basal		100						
	Pretest % Obj's Mastered		57.1	22.7		.31		
	Posttest %		75.3	17.3	77.7		6	5
	Reading Level*		2.8	1.2		.29		
Basal-Linguistic/Phonics		154						
	Pretest %		61.0	18.8		.25		
	Posttest %		80.0	17.1	79.3		3	2
	Reading Level		2.0	1.8		.25		
	Pre-Post Diff:		19.0					
Basal-Language Experience		143						
	Pretest %		61.3	19.3		.31		
	Posttest %		88.3	13.5	88.9		1	1
	Reading Level		3.7	1.5		.09		
	Pre-Post Diff:		27.0					
Basal-Programmed/Structured		50						
	Pretest %		59.0	17.8		.19		
	Posttest %		82.8	12.9	79.8		2	4
	Reading Level		4.4	1.7		.19		
	Pre-Post Diff:		13.1					
Linguistics/Phonics		26						
	Pretest %		57.1	16.3		.39		
	Posttest %		80.3	20.2	80.0		5	6
	Reading Level		4.8	1.9		.19		
	Pre-Post Diff:		23.2					
Linguistics/Phonics-Supplementary		30						
	Pretest %		60.7	13.2		.39		
	Posttest %		82.3	12.6	81.2		4	3
	Reading Level		4.1	2.6		-.09		
	Pre-Post Diff:		21.6					

*Reading Level is a coded variable, not the actual level.

+These numbers are based on logit transformation analysis; the rest of the data in the table are from the untransformed analysis.

Table XVII

Comparison of Reading Programs by Receiving Extra Help, Grade 1*

Program	Extra Help	Number of Students	Pretest %	SD	Posttest %(Unadj)	SD	Posttest % (Adj)	Pre-Post Diff(Unadj)	Reading Level	SD	Slopes Pre	Slopes RL
Basal	No	46	59.3	22.0	77.3	18.1	78.9	18.0	2.9	1.3	.31	.14
	Yes	54	55.1	23.3	73.7	16.5	76.6	18.6	2.7	1.2		
Basal-Linguistic/Phonics	No	78	66.4	19.9	84.6	16.9	82.3	20.2	4.4	1.8		
	Yes	76	55.6	15.9	75.2	16.1	77.2	19.6	3.3	1.7		
Basal-Programmed	No	24	72.6	14.2	85.8	10.5	80.9	13.2	5.1	1.3		
	Yes	26	65.8	20.3	80.1	14.5	79.1	14.3	3.7	1.8		
Basal-Language Experience	No	26	68.1	18.1	86.4	9.1	84.7	18.3	3.8	.9		
	Yes	117	59.8	19.3	88.7	14.4	89.0	28.9	3.7	1.6		
Lingustics/Phonics	No	16	58.9	17.7	80.4	23.3	79.7	21.5	4.8	1.5		
	Yes	10	54.2	14.0	80.1	15.1	79.1	25.9	4.9	2.5		
Linguistic/Phonics-Supplementary	No	14	55.8	22.0	85.0	10.5	86.8	29.2	3.4	.7		
	Yes	16	65.0	23.3	79.8	14.0	77.3	14.8	4.7	3.4		
Average		503 (Total)	60.9		81.6			20.2	3.7			

Note: Reading Level (RL) is a coded variable, not the actual reading level recorded by the teachers.
*Includes students in multi-level classes whose grade level averages 1.4 or less.

Table XVIII

Composition of Samples for Analyses of Reading Programs, Grade 1

Program	Number of Classes	Number of Students	Average Number of Students Per Class	Minimum Number of Students Per Class	Maximum Number of Students Per Class
Basal	12	100	8.33		
Basal-Linguistics/Phonics	22	154	7.00		
Basal-Programmed/Structured	6	50	8.33		
Basal-Language Experience	13	143	11.00		
Linguistics/Phonics	10	26	2.60		
Linguistics/Phonics-Supplementary	3	30	10.00		
Total number of Different Classes, Students	47	503		1	25

this variable did not seem to have a constant incremental effect on the average. We can see in Table XVII that in most reading program groups, but not all, the students who were receiving extra help in reading were those with the lower average pretest scores.

It should be noted in Table XVI that the regression coefficients and rank orders come from the analysis of covariance using percentage objectives mastered pre and post transformed to logits. The other data in the Table (as well as all data in Table XVII) are from the analysis using untransformed measures. The two analyses gave generally similar results. Thus, to avoid having to go through the process of reconverting the logits, we simply took percentages from the analyses using untransformed measures. The rank orders vary slightly. Therefore, there are times in which rank orders will be slightly different from the ranks that would be obtained from the percentages in the tables. All F-ratios are for the logit analyses. They are generally similar to those obtained from the analysis of untransformed data. One justification for using the transformation is that it eliminated two interactions in the one-way ANCOVAS, although it also introduced two that did not occur with the analyses of percentage data. Out of a total of 30 tests (all analyses), only 6 led to reversals of significance at the .05 level in comparing logit and percentage analyses of the same data.

There were several changes from the results of analyses of '75-'76 data. First, with respect to the list of types of programs analyzed, Basal Supplementary and Linguistics/Phoncis-Basal were not included in the '76-'77 analyses, while Linguistics/Phonics only and Linguistics/Phonics-Supplementary had not been included in the '75-'76 list.

Second, there was non-homogeneity of regression slopes for the one-way ANCOVA of the '75-'76 data, making the main effect not easily interpretable (the effect of type of program was not significant in '75-'76). For the

present data, the main effect is interpretable.

A further analysis of regions of significant differences between the regression slopes of the '75-'76 programs revealed that there were significant interactions between: Basal-Linguistics/Phonics and Basal-Supplementary; Basal-Linguistics/Phonics and Linguistics/Phonics-Basal; Basal-Language Experience and Basal-Supplementary; and Basal-Language Experience and Linguistics/Phonics-Basal. In each of these cases, the analyses revealed that students with above average pretest scores in the second type of program listed could be expected to do better on posttest than students with above average pretest scores in the first type of program. Significant interactions of this kind did not occur for the programs compared for the '76-'77 school year for Grade 1.

For the two-way ANCOVA's, there was a significant interaction of reading programs and whether or not receiving extra help in reading for the '76-'77 analyses, but not for the '75-'76 analyses. Thus, for the '76-'77 analyses, the effect of receiving extra help depend on the reading program. This was not the case for the '75-'76 analyses.

C. Comparisons of Reading Programs at the Second Grade Level

Eight types of program were selected for comparison at the 2nd grade level:

- Basal
- Basal-Linguistic Phonics
- Basal-Programmed/Structured
- Basal-Language Experience
- Basal-Individualized Reading
- Basal-Supplementary
- Linguistic/Phonics
- Linguistic/Phonics-Basal

Table XIX gives the descriptive statistics for these groups. The one-way ANOVA test of equality of group means was significant ($F_{7,604} = 4.031$, $p = .0003$). The ANCOVA test of adjusted posttest means was not significant ($F_{7,602} = 1.487$, $p = .1689$). The regression slopes of the groups were homogeneous ($F_{14,588} = 1.614$, $p = .0708$).

Table XX shows descriptive data for the two-way analysis of programs. In the two-way ANCOVA of whether or not the students were receiving extra help in reading, there was not a significant effect of reading programs ($F_{7,594} = .804$, $p = .584$), nor of extra help ($F_{1,584} = 1.641$, $p = .201$). The interaction was not significant ($F_{7,594} = .695$, $p = .677$). Table XXI gives sample composition information as in Table XVIII.

The results of the one-way ANCOVA lead to the conclusion that the types of reading programs compared did not differ significantly in effectiveness for the samples of programs and students analyzed. It should constantly be kept in mind that only students who took the same level of PRT pre and post were included in all the analyses reported here, and that of them, only students with valid reading program data, reading level data, and ID matched pre-post test data were included. These restrictions of course also applied to the '75-'76 data analyses.

There were several changes of results from '75-'76 to '76-'77. The '76-'77 comparisons included Basal-Language Experience, Basal-Individualized Reading, and Linguistics/Phonics-Basal programs. The '75-'76 comparisons had included Programmed/Structured-Basal programs. This type was not included in the '76-'77 comparisons.

In the '75-'76 analyses, results of the one-way ANCOVA led to the conclusion that there was a significant difference in effectiveness of the programs compared. For the '76-'77 analyses, the conclusion is that there were not significant differences in effectiveness of the samples of programs and students compared.

Table XIX

Statistical Data for Each Reading Program Compared, Grade 2

Program	Variable	Number of Students	Unadjusted Mean	Unadjusted SD	Adjusted Mean	Regression Slope+	Rank of:+ Unadj. M.	Rank of:+ Adj. M
Basal		93						
	Pretest % Obj's Mastered		56.6	22.4		.40		
	Posttest %		81.3	17.2	81.6		3	1
	Reading Level*		5.4	1.9		.29		
	Pre-Post Diff:		24.7					
Basal-Linguistic/Phonics		181						
	Pretest %		57.4	18.8		.68		
	Posttest %		79.8	16.2	79.7		4	4
	Reading Level		5.7	1.6		.11		
	Pre-Post Diff:		22.4					
Basal-Programmed/Structured		55						
	Pretest %		58.7	19.2		.41		
	Posttest %		76.3	20.5	77.5		7	7
	Reading Level		4.3	1.5		.11		
	Pre-Post Diff:		17.6					
Basal-Language Experience		61						
	Pretest %		57.8	20.4		.30		
	Posttest %		76.1	19.0	78.3		6	6
	Reading Level		4.2	1.9		.17		
	Pre-Post Diff:		15.9					
Basal-Individualized Reading		110						
	Pretest %		57.9	17.2		.51		
	Posttest %		80.1	14.3	80.2		5	3
	Reading Level		5.1	1.8		.12		
	Pre-Post Diff:		22.2					

Table XIX (Continued)

Program	Variable	Number of Students	Unadjusted Mean	Unadjusted SD	Adjusted Mean	Regression Slope+	Rank of:+ Unadj. M.	Rank of:+ Adj. M.
Basal-Supplementary		33						
	Pretest %		65.0	15.0		.14		
	Posttest %		87.1	7.9	79.8		1	5
	Reading Level		7.5	1.2		-.07		
	Pre-Post Diff:		22.1					
Linguistics/Phonics		39						
	Pretest %		49.4	16.8		.64		
	Posttest		70.2	18.2	76.0		8	8
	Reading Level		4.2	1.9		.08		
	Pre-Post Diff:		20.8					
Linguistics/Phonics-Basal		40						
	Pretest		59.9	19.5		.90		
	Posttest		81.8	18.9	79.6		2	2
	Reading Level		6.0	2.2		.06		
	Pre-Post Diff:		21.9					

*Reading Level is a coded variable, not the actual level.

+These numbers are from the logit transformation analysis; the rest of the data in the table are from the untransformed analysis.

Table XX

Comparison of Reading Programs by Receiving Extra Help, Grade 2*

Program	Extra Help	Number of Students	Pretest %	SD	Posttest %(Unadj)	SD	Posttest % (Adj)	Pre-Post Diff(Unadj)	Reading Level	SD	Slopes Pre	Slopes RL
Basal	No	27	65.6	18.8	81.2	14.8	78.0	15.6	5.2	1.1	.48	.15
	Yes	66	53.0	22.8	81.4	18.2	82.9	28.4	5.5	2.2		
Basal-Linguistic/Phonics	No	55	63.3	16.4	82.2	16.0	78.9	18.9	5.9	1.7		
	Yes	126	54.8	19.2	78.7	16.1	79.3	23.9	5.7	1.5		
Basal-Programmed/Structured	No	7	49.8	20.7	82.2	21.7	85.0	32.4		1.4		
	Yes	48	60.0	18.9	75.4	7.3	76.4	15.4		1.5		
Basal Individualized Reading	No	60	60.0	17.4	81.3	13.8	80.9	s1.3		1.9		
	Yes	50	55.4	16.8	78.6	14.8	79.7	23.2		1.7		
Basal/Language Experience	No	15	60.0	21.5	81.0	17.5	79.9	21.0		1.7		
	Yes	46	57.0	20.2	74.5	19.4	77.2	17.5		1.8		
Basal-Supplementary	No	9	71.7	11.6	89.2	6.4	78.9	17.5		.9		
	Yes	24	62.5	15.6	86.3	8.3	80.7	23.8		1.3		

Table XX (Continued)

Program	Extra Help	Number of Students	Pretest %	SD	Posttest %(Unadj)	SD	Posttest % (Adj)	Pre-Post Diff(Unadj)	Reading Level	SD	Slopes Pre	Slopes RL
Linguistic/Phonics	No	19	51.4	15.9	78.3	10.1	81.9	26.9		2.4		
	Yes	20	47.5	17.8	62.5	20.9	69.5	15.0		1.0		
Linguistics-Basal	No	23	69.4	18.3	88.2	21.8	80.5	18.8		2.1		
	Yes	17	47.1	12.7	73.1	13.7	78.6	26.0		1.5		
Average		612 (Total)	57.6		79.3			21.6	5.3			

Note: Reading Level (RL) is a coded variable, not the actual reading level recorded by the teachers.
*Includes students in multi-level classes whose grade level averages 1.5 to 2.4.

Table XXI

Composition of Samples for Analyses of Reading Programs, Grade 2

Program	Number of Classes	Number of Students	Average Number of Students Per Class	Minimum Number of Students Per Class	Maximum Number of Students Per Class
Basal	13	93	7.15		
Basal-Linguistics/Phonics	28	181	6.46		
Basal-Programmed/Structured	7	55	7.86		
Basal-Language Experience	10	61	6.10		
Basal-Individualized Reading	10	110	11.00		
Basal-Supplementary	7	33	4.71		
Linguistics/Phonics	7	39	5.57		
Linguistics/Phonics-Basal	5	40	8.00		
Total Number of Different Classes, students	62	612		1	21

The results of the two-way ANCOVA for the two years were different with respect to the conclusions about the effect of receiving extra help in reading. In '75-'76, it was concluded that students on the average who received extra help in reading gained significantly less than those who did receive extra help. That is, receiving extra help was not sufficient to enable those students to gain at the same rate as or better than those who did not receive extra help. From the '76-77 analyses it may be concluded that there was no significant effect associated with the variable. Students in given programs who received extra help tended to gain at the same rate as those who did not receive it. In the sense that the latter did not drop further behind the former, one can conclude there was positive benefit in receiving extra help for the '76-'77 samples analyzed, and that it did not depend on the program.

D. Comparisons of Reading Programs at the Third Grade Level

Six types of program were compared:

- Basal
- Basal-Linguistic/Phoncs
- Basal/Supplementary
- Basal-Individualized Reading
- Linguistics-Phonics

Descriptive statistics for the programs are given in Table XXII. The differences between the unadjusted means, as well as between the adjusted means, were highly significant (respectively: $F_{5,473} = 5.259$, $p = .0001$; $F_{5,471} = 3.290$, $p. = .0062$). The regression slopes of the groups, however, were not homogeneous ($F_{10,461} = 8.143$, $p = .0000$). Thus, the analysis of covariance is not directly interpretable.

Table XXIII gives descriptive statistics for the two-way analysis of covariance for third grade level programs. In the two-way ANCOVA, there was not a significant effect for reading programs ($F_{5,465} = 1.876$, $p = .097$), nor was

Table XXII

Statistical Data for Each Reading Program Compared, Grade 3

Program	Variable	Number of Students	Unadjusted Mean	Unadjusted SD	Adjusted Mean	Regression Slope+	Rank of:+ Unadj. M.	Rank of:+ Adj. M.
Basal		62						
	Pretest % Obj's Mastered		62.2	18.7		.79		
	Posttest %		79.4	14.7	90.3		5	3
	Reading Level*		7.0	2.4		.10		
	Pre-Post Diff:		17.2					
Basal-Linguistics		180						
	Pretest %		66.8	19.3		.54		
	Posttest %		81.7	17.3	79.5		.3	2
	Reading Level		7.7	1.9		.26		
	Pre-Post Diff:		14.9					
Basal-Language Experience		62						
	Pretest %		58.3	17.6		.26		
	Posttest %		69.6	20.0	79.6		6	4
	Reading Level		5.7	2.4		.13		
	Pre-Post Diff:		11.3					
Basal-Individualized Reading		61						
	Pretest %		68.0	18.8		.63		
	Posttest %		83.2	16.1	79.3		2	5
	Reading Level		8.3	1.1		-.15		
	Pre-Post Diff:		15.2					
Basal-Supplementary		33						
	Pretest %		64.6	24.0		.33		
	Posttest %		78.8	17.4	77.9		4	6
	Reading Level		7.7	1.7				
	Pre-Post Diff:		14.2					
Linguistics/Phonics		81						
	Pretest %		60.9	23.8		1.46		
	Posttest %		76.9	24.1	79.0		1	1
	Reading Level		7.1	2.3		.07		
	Pre-Post Diff:		16.0					

*Reading Level is a coded variable, not the actual level.

+These numbers are from the logit transformation analysis; the rest of the data in the table are from the untransformed analysis.

Table XXIII

Comparison of Reading Programs by Receiving Extra Help, Grade 3*

Program	Extra Help	Number of Students	Pretest %	SD	Posttest %(Unadj)	SD	Posttest % (Adj)	Pre-Post Diff(Unadj)	Reading Level	SD	Slopes Pre	Slopes RL
Basal	No	22	64.0	16.3	83.0	14.0	82.3	19.0	7.4	1.8	.62	.17
	Yes	40	61.2	20.0	77.4	19.0	79.3	16.2	6.9	2.7		
Basal-Linguistic/Phonics	No	55	62.5	21.3	77.8	20.6	77.1	15.3	7.7	1.8		
	Yes	125	68.6	18.1	83.5	15.4	80.1	14.9	7.7	1.8		
Basal-Language Experience	No	6	48.3	18.5	82.1	8.1	83.5	33.8	9.5	.5		
	Yes	56	48.3	17.7	68.3	20.4	79.1	20.0	5.3	2.1		
Basal-Individualized Reading	No	26	68.7	20.7	80.1	19.0	74.0	11.4	8.9	1.0		
	Yes	35	67.4	17.6	85.5	13.4	82.2	18.1	7.9	1.0		
Basal Supplementary	No	6	62.5	27.2	82.1	12.3	81.5	19.6	7.7	.8		
	Yes	27	65.0	23.7	78.1	18.5	76.4	13.1	7.7	1.8		
Linguistics/Phonics	No	50	67.8	22.3	85.7	18.0	83.8	17.9	7.3	2.2		
	Yes	31	49.7	21.5	62.8	26.1	69.6	13.1	6.8	2.5		
Averages		479	62.8		79.0			17.7	7.3			

Note: Reading Level (RL) is a coded variable, not the actual reading level recorded by the teachers.
*Includes students in multi-level classes whose grade level averages 2.5 to 3.4.

Table XXIV

Compositions of Samples for Analyses of Reading Programs, Grade 3

Program	Number of Classes	Number of Students	Average Number of Students Per Class	Minimum Number of Students Per Class	Maximum Number of Students Per Class
Basal	14	62	4.43		
Basal-Linguistics/Phonics	26	180	6.92		
Basal-Language Experience	7	62	8.86		
Basal-Individualized Reading	12	61	5.08		
Basal-Supplementary	10	33	3.30		
Linguistics/Phonics	12	81	6.75		
Total Number of Different Classes, Students	47	479		1	21

there a significant effect for receiving extra help in reading ($F_{1,465} = .450$, $p = .503$). The interaction was significant ($F_{5,465} = 5.908$, $p = .000$). Table XXIV gives the sample compositions for Grade 3.

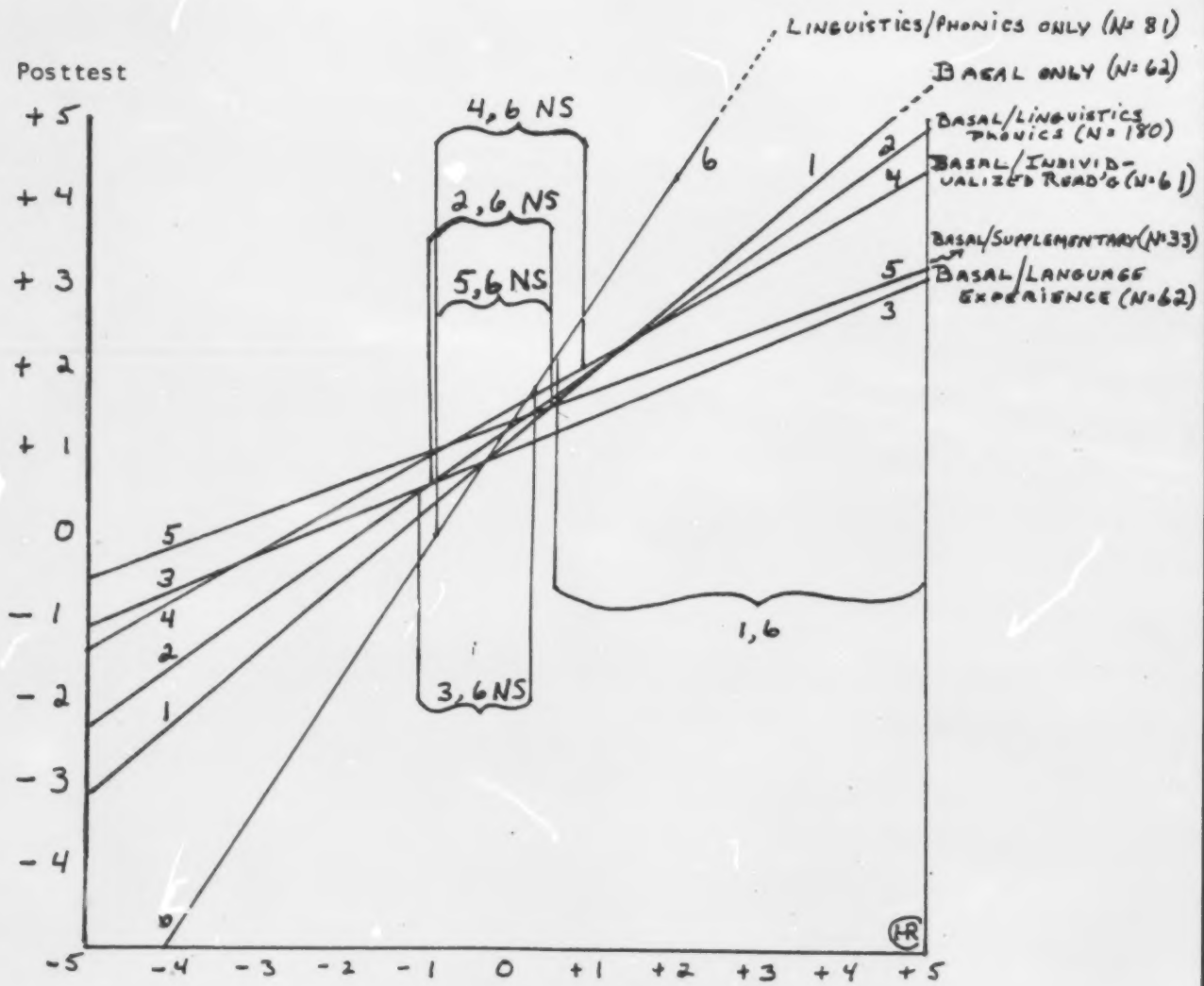
The analyses suggest that there were differences in the effectiveness of different types of programs, but that the relationships were complex. That is, different programs seemed to be affecting different groups differently. Figure 1 shows the regression lines of each program for posttest regressed on pretest. The axes are for the logit scale. Figure 1 also shows the regions of significant differences between slopes for designated pairs of types of reading programs. Figure 1 shows graphically what is meant by a complex relationship of differences among programs, or by an interaction.

The one-way ANCOVA leads to the conclusion that there was a significant difference in effectiveness among the types of reading programs compared, but that the difference was complex (there was non-homogeneity of regression slopes). Using the logit transformed data, Neymann-Johnson regions of significance (at the .05 level) between regression slopes were computed for each pair of type of program. Figure 1 shows clearly that it is the Linguistic-Phonics type that produced significant interactions with the other types of programs insofar as regression of posttest on pretest is concerned (analyses of regressions of posttest on reading level of most used material were not undertaken for this report). Since it will be found that similar analyses, figures, and interpretations will be given for Grades 4 and 6, it is worth examining Figure 1 in some detail.

First, the scales for pretest and for posttest pertain to the logit transformation and cover only the approximate range of possible values for it. A logit value of 0 is equivalent to 50% objectives mastered (either pre or post). A logit of around +5 or -5 is equivalent to 100% and 0% objectives mastered, respectively. We say "around" because the exact logit values for those percents

Figure 1

Regions of Significant Difference Between Pairs of Reading Programs, Grade 3



Note:

Brackets designate regions of significance for the indicated pairs of regression slopes 6/reading programs. Regions labelled NS are regions in which the difference between slopes is not significant.

depend on the test level, since the logit is based on the number of objectives and different test levels have different numbers of objectives. A logit of +1 or -1 is approximately equivalent to 75% or 25% respectively. A logit of +2 or -2 is approximately equivalent to 90% or 10% respectively, and a logit of +2.5 is roughly equal to about 92%. Thus, in Figure 1, the fact that the regression line for Linguistic/Phonics reaches a maximum positive value of +2.5 on the pretest scale means that there were no students in the sample for that type of program who had a pretest score of above about 92% reading objectives mastered. The same, of course, might have been true for the other types of programs, but the linear projections of the regression lines did not exceed +5 on the posttest scale (except, as can be seen, for the Basal only type). The regression tests and regression lines are based on the assumption of linear regression. Slopes were plotted simply by fairing lines through three points whose values were determined from the regression equation for each program regardless of the actual distribution of pre and posttest scores.

For further discussion for the moment, let us refer to types of programs by the numbers shown in Figure 1 as a matter of convenience. Note that the brackets in fact demarcate two types of regions: 1) a region or zone along the pretest axis in which the regression lines for a pair of programs differ significantly at the .05 level; and 2) a region or zone within which the slopes of a pair of programs do not differ significantly (designated as NS in Figure 1) from each other at the .05 level. The brackets for the comparison of programs 1 and 6 are of the first type; the brackets for the comparisons of the other pairs shown are of the second type. Thus, for programs 1 and 6, there is a significant difference between regression lines for pretest values of about +0.6 or better, out to the end of the pretest range (+5.0). The two lines do not differ significantly over the rest of the range of the pretest scale (i.e., from less than +.6 to about -4.1). Actually, as we can see in Figure 1,

and as we have stated before, the region of significance can only extend out to about +2.5 on the pretest range since that is where the program 6 line reaches the maximum value possible (+5.0) on the posttest scale. We also, of course, cannot tell from Figure 1 whether there were any students in the sample for program 1 who had any pretest scores of greater than +2.5. The vast majority of students in all programs in fact fall between ±2.5, however.

Now look at the rest of the regions bracketed in Figure 1. They are of the second type, having boundaries around the regions along the pretest scale within which the slopes of the pairs of programs designated did not differ significantly at the .05 level of significance. Regions along the pretest scale outside the brackets in either direction are regions in which the pairs of programs indicated do differ significantly (out, of course, to about the limits of the line for program 6).

Note that for most of the pairs of programs in Figure 1, no bracket of either type is shown. If no bracket is shown for a pair of programs (e.g., 1, 2 in Figure 1), it means that there was no significant difference between the slopes of those two programs over the whole possible range of the pretest variable as measured in logits. This is the case for most of the pairs of programs in Figure 1. Of the (6 x 5)/2 pairs, only 5 have regions of significance brackets, and all 5 of those pairs involve program 6! We can now see at a glance why we earlier identified program 6 (Linguistic/Phonics only) as the type of program that caused the test of homogeneity of regression to be significant (that is, as the cause of the significant interaction). As a general note, if the test of differences between means had been significant, but the test for homogeneity or equality of slopes not significant, the lines would appear separated but essentially parallel.

Before interpreting the results shown in Figure 1, let us examine one more important fact in it. For the convenience of the reader, the program

numbers have been written on the appropriate regression lines at both ends of the pretest (or X-) axis. Note that at the lower (left) end of the pretest scale, program 6 is on the bottom while at the upper (right) end it is at the top. The regression line of program 6 crosses the lines of the other programs, roughly around 0 (or 50% objectives mastered) on the pretest scale. When two regression lines cross within the range of the pretest and posttest variables, they are called <u>disordinal</u>. Thus, the regression slopes of program 6 and each of the other programs in Figure 1 are disordinal. Note, incidentally, that the lines for programs 3 and 5 are not disordinal; they do not cross within the range of the pre and posttest variables. The line for program 5 is higher (though not significantly so) than that for program 3 over the whole range.

The significant disordinal relationships in Figure 1 tell us something very important. They tell us that, on the average, students in program 6 who did poorly on the pretest did <u>poorer</u> than expected on posttest compared to similar students in the other types of programs, while students in program 6 who did well on pretest on the average did <u>better</u> than expected compared to similar students in the other types of programs (only the latter part of this conclusion applies to the 1, 6 comparison, since those slopes are not significantly different from about +.6 on down to the bottom of the pretest range). This of course is precisely what the term "interaction" means as applied to these analyses of program effectiveness. The regions of significance (or the bracketed regions of non-significance) also tell us that students in program 6 who scored around the middle of the pretest range on pretest on the average did no better or worse on posttest than similar students in the other programs. For such students it made no difference what type of reading program they were in; all six types of programs were equal in effectiveness for the samples of students involved in the analysis. It is important in any case to note that

when we say "better than expected" or "poorer than expected," we mean in relation to the performance of the specific samples of students in the particular types of programs used in this analysis. We do not mean in relation to an external standard. The comparisons are of the programs and samples of students with each other. That is why final conclusions about the relative effectiveness of types of reading programs must be qualified. If the same relationships hold for analyses of the '77-'78 data, far more implication of absolute effectiveness can be drawn.

We can now draw some simple conclusions and implications from Figure 1. A straight linguistics/phonics type of reading program, compared to the other types appears to work particularly well, on the average, for students who mastered around 67%-82% of the reading objectives or more on the pretest (Fall '76 testing). That type of reading program appeared to work very poorly (in terms of posttest achievement), on the average, for students who showed mastery of about 25% of the reading objectives or less on the pretest. The exception to this latter statement was for those students in the linguistic/phonics only type of program (#6) and in the basal only type of program (#1). Students in those two types who mastered less than about 85% of the reading objectives on the pretest did comparably on the posttest. Otherwise, students in all six programs who scored around 25-67% on pretest did about equally well on the average on posttest. If one ignores the pure linguistics/phonics type of program (#6) altogether, the other types of programs were equally effective, on the average, regardless of how students did on pretest.

It is, of course, entirely possible that the differences in effectiveness that were observed were caused by factors other than the type of reading program. We will analyze other factors in future analyses. At this point the differences in effectiveness can be said to be associated with type of reading program, but by no means necessarily caused by it.

It is not possible to tell from the '75-'76 analyses whether the present results demonstrate a continuing pattern insofar as Linguistic/Phonics type of program is concerned because the sample size was inadequate to include that type in the '75-'76 analysis of third grade programs. There was, however, a significant disordinal relationship in the '75-'76 analysis of third grade programs between the Basal only and Basal-Linguistics/Phonics programs. The relationship of the regression lines of those two types in '75-'76 was precisely like that of the two (1,2) in Figure 1 for the '76-'77 data, although in '76-'77 the separation between lines from the middle of the pretest range on down is not quite sufficient (given the sample sizes) to be statistically significant. The fact that the disordinal pattern is the same for the two years, however, is important. It is suggested that third-grade teachers contemplating a choice of type of reading program (a pure basal type, or basal supplemented with linguistics/phonics materials) for students who score below average on the fall PRT should give strong consideration to the basal-linguistics/phonics type in preference to just a basal program. It should be remembered, however, that until an analysis of all students is completed, this suggestion pertains only to low scoring students who are likely to be given the same PRT test level in the following spring.

The reader may wonder why the same suggestion should not be made with respect to other pairs of programs involving the basal only type shown in Figure 1 (e.g., 1, 4; 1, 3; 1, 5). The reason that no particular suggestion or recommendation is made about choosing among them is that while those types were compared both with the '75-'76 data and the '76-'77 data, in neither evaluation was a significant difference between them found, either overall or over a particular pretest region. Thus, while the patterns of regression lines as such may suggest an advantage of one or more types of programs (e.g., 3 or 4 or 5) over a pure basal program (1), in <u>neither</u> year was there statistical support

for deciding in favor of one over the other, given the samples and sample sizes available for analysis. It is possible that with larger sample sizes to provide more powerful tests, such support will be forthcoming. As a general decision rule for drawing conclusions about the relative effectiveness of types of programs, it is proposed that no conclusion be drawn, even tentatively, unless the same performance relationship of two or more programs can be shown over at least two years, and unless in at least one of those years a statistically significant difference can be shown.

The reader may also wonder whether the results depicted in Figure 1 for particular pairs of programs were not simply due to the work of two or three teachers for each type of program. Reference to Table XXIV will show that the samples of students for each program came from an average of over 13 different classes, and that in no case were less than 7 different classes involved.

For the present data ('76-'77), the results of the two-way ANCOVA suggest that, as in the first grade results, the effectiveness of receiving extra help in reading depends on the type of reading program (that is, there is an interaction). Since that is a different result from that obtained from analysis of the '75-'76 third grade programs, no clear conclusion can be drawn about that program component as yet.

E. Comparisons of Reading Programs at the Fourth Grade Level

Six types of program met the selection criteria:

- Basal
- Basal-Linguistic/Phonics
- Basal-Individualized Reading
- Basal-Supplementary
- Linguistics-Phonics
- Linguistics-Phonics/Individualized Reading

Table XXV presents the descriptive statistics for each type of program. The test of differences of unadjusted means was significant, as was the test of adjusted means at the .05 level (respectively: $F_{5,742} = 8.223$, $p = .0000$; $F_{5,740} = 5.5566$, $p = .0000$). The regression slopes were not homogeneous ($F_{10,730} = 5.4356$, $p = .0000$).

Table XXVI gives descriptive statistics for the two-way analysis of covariance. In this analysis, the reading program effect was significant ($F_{5,734} = 3.199$, $p = .007$); The test of receiving extra help in reading was not significant ($F_{1,734} = 2.354$, $p = .125$). The interaction was not significant ($F_{5,734} = .490$, $p - .784$). The regression coefficient for reading level of materials was not significantly different from zero. Table XXVII gives the sample compositions of the different programs.

From the one-way ANCOVA it may be concluded that there appears to be a difference in effectiveness of the reading programs, but that again relationships are complex. Figure 2 shows the regression lines of the group of types of program. Figure 2 depicts the same type of information as Figure 1 in Section VI D, except that in Figure 2 bars instead of brackets are used to demarcate regions of significance to avoid undue cluttering of the chart. The principles for interpreting Figure 2 are the same as those for Figure 1.

It may be seen in Figure 2 that two types of reading programs produced the significant interactions: Linguistic/Phonics (5), and Linguistic/Phonics-Individualized Reading (6). For students who on pretest showed mastery of about 50% of the reading objectives or more on the test level administered, a pure linguistics/phonics type of reading program was significantly more effective on the average than any of the other 5 types analyzed. On the other hand, for students who did about 70% or better on the pretest (around ± 7.8 in Figure 2), a linguistics/phonics program supplemented with individualized reading materials was associated with significantly less effectiveness on

Table XXV

Statistical Data for Each Reading Program Compared, Grade 4

Program	Variable	Number of Students	Unadjusted Mean	Unadjusted SD	Adjusted Mean	Regression Slope+	Rank of:+ Unadj. M	Rank of:+ Adj. M
Basal		128						
	Pretest & Obj's Mastered		55.2	22.0		.75		
	Posttest %		68.3	19.2	69.5		5	5
	Reading Level*		9.0	2.3		-.03		
	Pre-Post Diff:		13.1					
Basal-Linguistic/Phonics		219						
	Pretest %		56.1	18.3		.58		
	Posttest		70.0	16.9	72.0		3	3
	Reading Level		8.2	2.4		.09		
	Pre-Post Diff:		13.9					
Basal-Individualized Reading		143						
	Pretest %		54.6	16.5		.88		
	Posttest		65.7	17.1	71.4		2	2
	Reading Level		8.7	1.7		-.07		
	Pre-Post Diff:		11.1					
Basal-Supplementary		167						
	Pretest %		68.4	20.9		.78		
	Posttest		55.9	19.7	70.0		4	4
	Reading Level		8.7	1.9		.01		
	Pre-Post Diff:		12.5					
Linguistics/Phonics		46						
	Pretest %		81.4	17.3		.84		
	Posttest %		62.4	15.1	78.7		1	1
	Reading Level		7.9	1.9		-.21		
	Pre-Post Diff:		19.0					
Linguistics/Phonics-Indiv. Rdg.		45						
	Pretest %		55.4	22.7		.34		
	Posttest %		64.9	18.7	66.5		6	6
	Reading Level		9.3	1.6		.11		
	Pre-Post Diff		9.5					

*Reading Level is a coded variable, not the actual level.

+The numbers come from the logit transformation analyses; the rest of the data in the table are from untransformed analyses.

Table XXVI

Comparison of Reading Programs by Receiving Extra Help, Grade 4*

Program	Extra Help	Number of Students	Pretest %	SD	Posttest %(Unadj)	SD	Posttest % (Adj)	Pre-Post Diff(Unadj)	Reading Level	SD	Slopes Pre	Slopes RL
Basal	No	47	53.6	22.7	67.9	20.4	69.7	14.3	10.4	1.5	.67	.03
	Yes	81	56.1	21.7	68.6	18.6	69.9	12.5	8.2	2.4		
Basal-Linguistics/Phonics	No	122	57.7	17.3	71.3	15.7	71.2	13.6	9.0	2.3		
	Yes	97	54.8	18.9	68.9	18.0	71.4	14.1	7.6	2.3		
Basal-Individualized Reading	No	90	64.7	17.9	75.5	16.3	70.9	10.8	9.2	1.4		
	Yes	53	64.5	14.0	75.9	18.3	72.1	11.4	8.1	2.0		
Basal Supplementary	No	97	55.3	21.8	67.1	19.7	68.7	11.8	8.8	1.9		
	Yes	70	56.7	19.6	70.1	19.8	70.9	13.4	8.5	1.9		
Linguistics/Phonics	No	11	53.9	17.2	74.4	15.5	76.6	20.5	9.0	1.2		
	Yes	35	65.1	16.7	83.6	14.5	79.6	18.5	7.5	2.0		
Linguistics	No	30	62.8	17.6	70.4	14.1	66.8	7.6	9.6	1.4		
	Yes	15	40.6	25.0	54.0	22.3	64.7	13.4	8.9	1.9		
Average		748 (Total)	57.9		70.8			13.5	8.6			

Note: Reading Level (RL) is a coded variable, not the actual reading level recorded by the teachers.
*Includes students in multi-level classes whose grade level averages 3.5-4.4.

Table XXVII

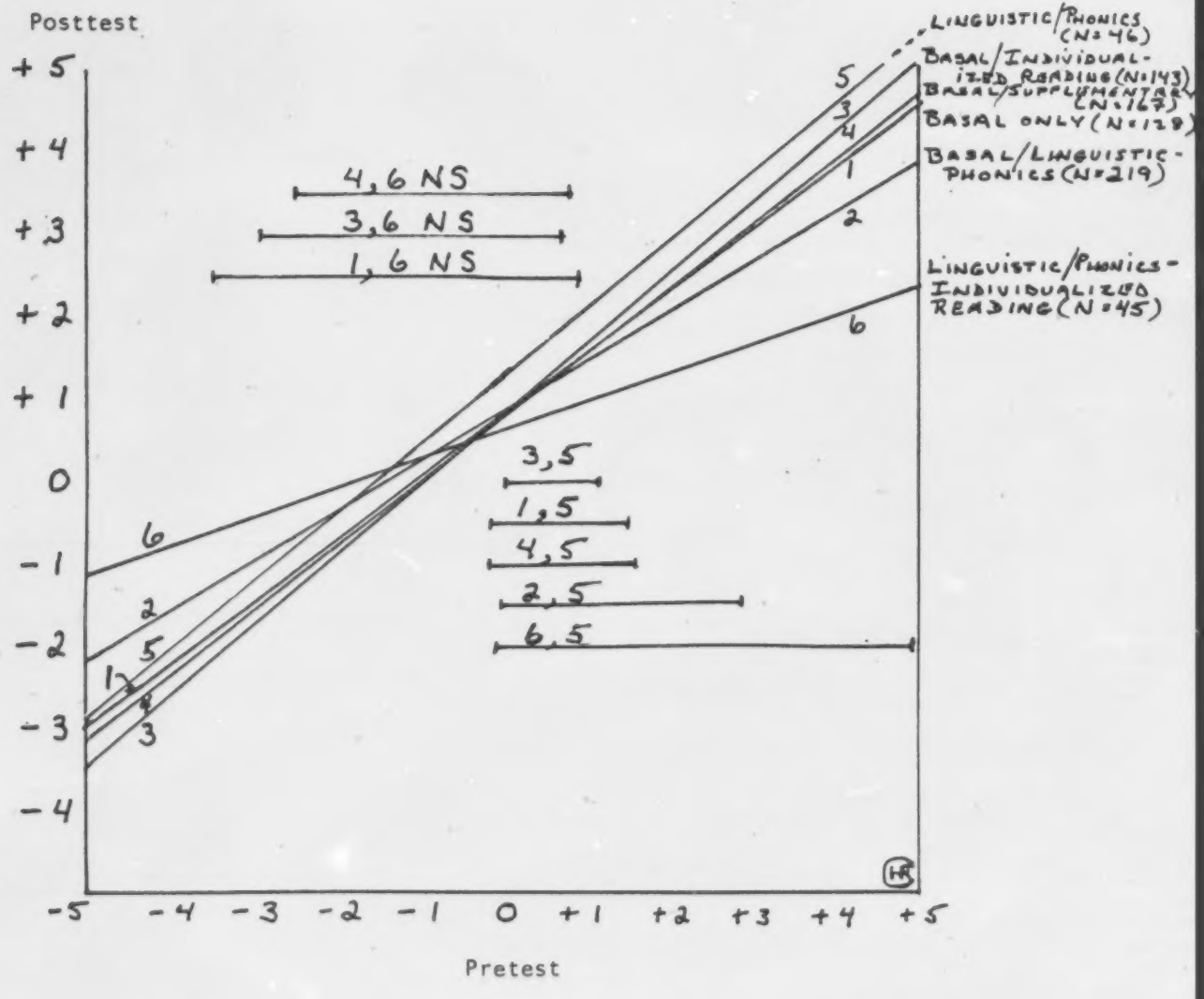
Composition of Samples for Analyses of Reading Programs, Grade 4

Program	Number of Classes	Number of Students	Average Number of Students Per Class	Minimum Number of Students Per Class	Maximum Number of Students Per Class
Basal	25	128	5.12		
Basal-Linguistics/Phonics	27	219	8.11		
Basal-Individualized Reading	21	143	6.81		
Basal-Supplementary	22	167	7.59		
Linguistics/Phonics	7	46	6.57		
Linguistics/Phonics-Individualized Reading	6	45	7.5		
Total Number of Different Classes, Students	72	748		1	34

143

Figure 2

Regions of Significant Difference Between Pairs of Reading Programs, Grade 4



Note:

Because of the large number of interactions, bars are used to depict regions of significant differences or of non-significance (NS) between the regression slopes of reading programs. Numbers designate the pairs of programs to which the bars pertain.

the average than 4 of the other 5 types of programs. Otherwise, programs were essentially comparable in effectiveness (i.e., not statistically significantly different) on the average. Note, however, that with respect to program 5 in comparison to programs 1, 3 and 4, the effectiveness of the pure linguistic/phonics type of program was comparable to that of the other three types, on the average, for students who scored substantially above average on pretest (roughly 77-87% or better). Thus, with respect to a choice of program 5 in relation to 1 or 3 or 4 (see Figure 2), program 5 was significantly more effective with students who were somewhat below to somewhat above average on pretest, but not significantly more or less effective with students who scored considerably above or below average.

In the analyses of '75-'76 programs, four types were compared: Basal, Basal-Linguistic/Phonics, Basal-Individualized Reading, and Basal-Supplementary. The results of the one-way ANCOVA were that the difference between means nearly reached significance ($p = .057$), and that the slopes could be considered homogeneous (no significant interactions). Since the linguistic/phonics-individualized reading type of program was not included in the analysis for lack of an adequate sample size, a conclusion leading to a recommendation based on data for two years cannot be drawn (see decision rule, Section VI D). It will be necessary to obtain the results of analyses of the '77-'78 data to see if a reliable pattern of performance is discernible.

The results of the two-way ANCOVA were the same as for the '75-'76 data. The effect of reading program was significant; the effect of receiving extra help in reading was not. There was not a significant interaction between the two variables. In effect, when type of program was taken into account, and when adjustment was made for pretest and level of most frequently used reading material, students who received extra help in reading seemed to gain at about the same rate as those who did not. The gap did not close, but neither did it tend to widen.

F. Comparisons of Reading Programs at the Fifth Grade Level

Six types of reading program were compared:

- Basal only
- Basal-Linguistic/Phonics
- Basal-Language Experience
- Basal-Individualized Reading
- Basal-Supplementary
- Linguistic/Phonics only

Table XXVIII gives the descriptive statistics for each program. Both the one-way test of unadjusted and of adjusted posttest means were significant at the .05 level (respectively: $F_{5,844} = 2.741$, $p = .0185$; $F_{5,584} = 4.006$, $p = .0014$). Regression slopes were equal ($F_{10,574} = 1.242$, $p = .2612$).

Descriptive statistics for the two-way analyses of covariance of groups are given in Table XXIX. In this analysis also there was a significant effect for reading programs ($F_{5,578} = 3.433$, $p = .005$), but not for receiving extra help ($F_{1,578} = .044$, $p = .835$). The interaction was significant ($F_{5,578} = 2.655$, $p = .022$). In this sample, the regression coefficient for reading level of materials was not significantly different from zero. Table XXX shows the sample compositions for the programs compared.

There is no need for a figure showing the regression lines of these programs. They are essentially parallel, as the results of the test of homogeneity of regression indicates. The information about effectiveness is contained in the separation between the lines, which is entirely measured by the differences between means as long as linearity is assumed. Thus, an interpretation of results of the one-way ANCOVA can be made simply by referring to Table XXVIII From t-tests of differences between adjusted means it was found that the pure Basal type of program was significantly less effective than: Basal-Linguistic/Phonics; Basal-Language Experience; and Basal-Individualized Reading. The

Table XXVIII

Statistical Data for Each Reading Program Compared, Grade 5

Program	Variable	Number of Students	Unadjusted Mean	Unadjusted SD	Adjusted Mean	Regression Slope+	Rank of:+ Unadj. M.	Rank of:+ Adj. M.
Basal		112						
	Pretest % Obj's Mastered		62.2	16.3		.69		
	Posttest %		69.1	15.2	68.9		6	5
	Reading Level*		9.7	1.9		-.03		
	Pre-Post Diff:		6.9					
Basal-Linguistics/Phonics		254						
	Pretest %		59.1	19.3		.66		
	Posttest %		70.7	17.9	72.0		4	1
	Reading Level		9.6	2.2		.03		
	Pre-Post Diff:		11.6					
Basal-Language Experience		39						
	Pretest %		63.6	17.0		.92		
	Posttest %		73.4	16.3	72.2		2	2
	Reading Level		8.2	1.9		.01		
	Pre-Post Diff:		9.8					
Basal-Individualized Reading		101						
	Pretest %		62.0	19.8		.74		
	Posttest %		74.6	17.6	73.2		1	4
	Reading Level		10.0	1.2		.16		
	Pre-Post Diff:		12.6					
Basal-Supplementary		54						
	Pretest %		58.8	19.2		.63		
	Posttest %		70.2	16.9	71.7		5	3
	Reading Level		9.2	2.2		.04		
	Pre-Post Diff:		11.4					
Linguistic/Phonics-Individualized Reading		32						
	Pretest %		68.0	13.6		.95		
	Posttest %		73.7	15.0	68.6		3	6
	Reading Level		10.8	.8				
	Pre-Post Diff:		5.7					

*Reading Level is a coded variable, not the actual level.

+These numbers come from analyses of logit transformations; the rest of the data in the table are from untransformed analyses.

Table XXIX

Comparison of Reading Programs by Receiving Extra Help, Grade 5*

Program	Extra Help	Number of Students	Pretest %	SD	Posttest %(Unadj)	SD	Posttest % (Adj)	Pre-Post Diff(Unadj)	Reading Level	SD	Slopes Pre	Slopes RL
Basal	No	49	64.6	16.6	68.7	16.3	66.0	4.1	10.3	1.1		
	Yes	63	60.3	18.9	69.4	14.3	70.0	6.4	9.2	2.2		
Basal-Linguistics/Phonics	No	86	57.5	19.8	70.3	18.2	72.5	12.8	9.4	2.3		
	Yes	168	59.9	19.0	71.0	17.9	71.6	11.1	9.6	2.1		
Basal-Language Experience	No	11	63.9	19.9	76.1	12.1	75.4	12.2	7.7	1.8		
	Yes	28	63.5	16.1	72.3	17.8	71.5	8.8	8.4	1.9		
Basal-Individualized Reading	No	61	65.1	18.1	78.8	14.7	75.9	13.7	10.1	.8		
	Yes	40	57.4	21.5	68.2	19.9	70.3	10.8	9.7	1.6		
Basal-Supplementary	No	26	67.4	16.7	74.7	16.0	71.2	7.3	8.7	2.7		
	Yes	28	50.9	18.2	66.1	16.9	72.3	15.2	9.7	1.4		
Linguistics-Individualized Reading	No	10	67.3	15.1	70.0	18.2	65.4	2.7	10.7	.9		
	Yes	22	68.3	13.3	75.4	13.5	70.1	7.1	10.8	.7		
Averages		592 (Total)	60.9		71.4			9.4	9.6			

Note: Reading Level (RL) is a coded variable, not the actual reading level recorded by the teachers.
*Includes students in multi-level classes whose grade level averages 4.5-5.4.

Table XXX

Composition of Samples for Analyses of Reading Programs, Grade 5

Program	Number of Classes	Number of Students	Average Number of Students Per Class	Minimum Number of Students Per Class	Maximum Number of Students Per Class
Basal	17	112	6.59		
Basal-Linguistic/Phonics	32	254	7.94		
Basal-Language Experience	6	39	6.50		
Basal-Individualized Reading	10	101	10.1		
Basal-Supplementary	10	54	5.40		
Linguistic/Phonics	6	32	5.33		
Total Number of Different Classes, Students	59	592		1	23

pure Linguistic/Phonics type was significantly less effective than the Basal-Individualizing Reading type of program. And the Basal-Linguistic/Phonics type was associated with significantly more effectiveness than the Basal-Individualized Reading type, on the average. Succinctly the 5th ranked program (see rank orders for adjusted means, Table XXX) was less effective than the programs ranked 1, 2, and 4. The 6th ranked program was less effective than the 4th ranked program. And the 1st ranked program was more effective than the 4th ranked program by the methods of measurement and analysis used in this study. Although the seeming disparity in relationships thus found (e.g., why should not the 5th ranked type of program also have been found to be more effective than the 3rd ranked one?) can not be explained with certainty, a reasonable hypothesis is that the disparities are the result of the very small sample sizes of three of the types of program.

In the '75'76 analyses, there was not a significant difference between adjusted means, although there was a significant interaction. Thus, results with the '76-'77 data indicate some changes from the preceding year. There are some commonalities, however, if one looks only at adjusted means for the two years. In '75-'76, the Basal only type of program ranked lowest (6)' in '76-'77 it continued to rank low (5th out of 6). In both years, sample sizes for that type were over 100, so some confidence may be placed in the result. The Basal-Language Experience type ranked second both years, and the Basal-Individualized Reading type ranked 4th both years. It was the Basal-Language Experience type of program, however, that produced the significant interaction in '75-'76. Thus, it is not yet possible to say with assurance whether this type is really disordinal with respect to other types. The Basal-Lingustic/Phonics type of program showed a major reversal from 5th (out of 6) in '75-'76 to 1st in '76-'77.

The '76-'77 two-way ANCOVA results also showed some changes from '75-'76. While in both years the effect of receiving extra help in reading was not

statistically significant, in the present analysis ('76-'77) there was a significant interaction of that variable with reading program, suggesting again that the effect of extra help depended on the type of reading program.

G. Comparisons of Reading Programs at the Sixth Grade Level

Seven types of reading program were compared:

- Basal only materials
- Basal-Linguistics/Phonics materials
- Basal-Individualized Reading materials
- Basal-Supplementary materials
- Linguistics/Phonics materials
- Linguistics/Phonics-Supplementary materials
- Supplementary

Table XXXI summarizes descriptive statistics for each program. The test of differences of unadjusted means for the one-way analysis of variance was significant, as was the one-way ANCOVA test of adjusted means ($F_{6,831} = 2.550$, $p = .0187$; $F_{6,829} = 3.569$, $p = .0017$). Regression slopes were heterogeneous ($F_{12,817} = 3.321$, $p = .0001$).

Table XXXII gives descriptive statistics for the two-way ANCOVA. The variable, reading program, was significant $F_{6,822} = 2.840$, $p = .010$); the variable, extra help, was not $F_{1,862} = 1.587$, $p = .208$). The interaction was not significant ($F_{6,922} = .977$, $p = .439$).

Table XXXIII gives the sample compositions for programs in these analyses. Figure 3 shows the regression lines of posttest regressed on pretest for each program in the analysis.

It can be seen at a glance in Figure 3 that it was the Basal-Supplementary type of program (4) that produced the significant interaction in the one-way ANCOVA. It can also be seen that the regression line for that type is disordinal

Table XXXI

Statistical Data for Each Reading Program Compared, Grade 6

Program	Variable	Number of Students	Unadjusted Mean	Unadjusted SD	Adjusted Mean	Regression Mean	Rank of:+ Unadj. M.	Rank of:+ Adj. M.
Basal		92						
	Pretest % Obj's Mastered		63.7	18.6		.78		
	Posttest %		72.0	17.0	76.8		2	3
	Reading Level*		10.2	2.1		-.04		
	Pre-Post Diff:		8.3					
Basal-Linguistic/Phonics		368						
	Pretest %		60.1	21.1		.73		
	Posttest %		71.9	18.8	72.1		3	2
	Reading Level		10.7	1.7		.05		
	Pre-Post Diff:		11.8					
Basal-Individualized Reading		72						
	Pretest %		58.9	21.4		.76		
	Posttest %		68.6	19.9	69.2		4	7
	Reading Level		.8	1.2		.05		
	Pre-Post Diff:		9.7					
Basal-Supplementary		134						
	Pretest %		62.7	18.4		.27		
	Posttest %		72.3	17.3	70.9		1	1
	Reading Level		10.8	1.6		.05		
	Pre-Post Diff:		9.6					
Linguistics/Phonics		77						
	Pretest %		56.6	23.5		.61		
	Posttest %		66.7	19.8	68.1		6	5
	Reading Level		11.0	1.5				
	Pre-Post Diff:		10.1					

Table XXXI (Continued)

Program	Variable	Number of Students	Unadjusted Mean	Unadjusted SD	Adjusted Mean	Regression Slope+	Rank of:+ Unadj. M.	Rank of:+ Adj. M.
Linguistics/Supplementary		30						
	Pretest %		57.0	20.2		.89		
	Posttest %		66.1	22.5	67.9		7	6
	Reading Level		11.6	2.0		-.06		
	Pre-Post Diff:		9.1					
Supplementary		65						
	Pretest %		59.3	25.7		.65		
	Posttest %		66.3	24.3	66.2		5	4
	Reading Level		11.2	1.4		.13		
	Pre-Post Diff:		7.0					

*Reading Level is a coded variable, not the actual level.

+These numbers are from the analyses of logit transformations; the rest of the data in the table are from untransformed analyses.

Table XXXII

Comparison of Reading Programs by Receiving Extra Help, Grade 6*

Program	Extra Help	Number of Students	Pretest %	SD	Posttest %(Unadj)	SD	Posttest % (Adj)	Pre-Post Diff(Unadj)	Reading Level	SD	Slopes Pre	Slopes RL
Basal	No	54	70.2	13.7	76.6	11.9	70.9	6.4	10.4	1.6		
	Yes	38	54.3	20.7	65.5	20.9	69.8	11.2	9.9	2.7		
Basal-Linguistics/ Phonics	No	202	64.4	20.5	75.6	18.6	73.0	11.2	10.9	1.4		
	Yes	166	54.9	20.7	67.3	18.0	70.8	12.4	10.5	2.0		
Basal-Individualized Reading	No	47	64.3	18.5	72.9	18.3	70.4	8.6		1.0		
	Yes	25	48.6	22.9	60.3	20.6	67.6	11.7	10.5	1.5		
Basal-Supplementary	No	80	62.4	19.8	72.5	17.2	71.0	10.1	11.1	1.1		
	Yes	54	63.2	16.2	72.0	17.7	70.6	8.8	10.3	2.0		
Linguistics/ Phonics	No	59	60.5	20.9	70.6	16.6	70.0	10.1	11.4	.9		
	Yes	18	43.6	27.5	53.7	24.0	64.6	10.1	9.6	2.2		
Linguistics/ Supplementary	No	24	60.5	17.6	67.7	21.5	66.6	7.2	12.1	.3		
	Yes	6	43.3	25.6	59.4	27.4	70.3	16.1	9.8	4.2		
Linguistics	No	24	65.6	25.7	73.1	24.2	69.6	7.5	2.4	2.1		
	Yes	41	55.6	25.3	62.5	23.8	65.0	6.9	4.1	.9		
Average		838 (Total)	60.3		70.5			9.9	10.8			

Note: Reading Level (RL) is a coded variable, not the actual reading level recorded by the teachers.
*Includes Students in multi-level classes whose grade level averages 5.5-6.4.

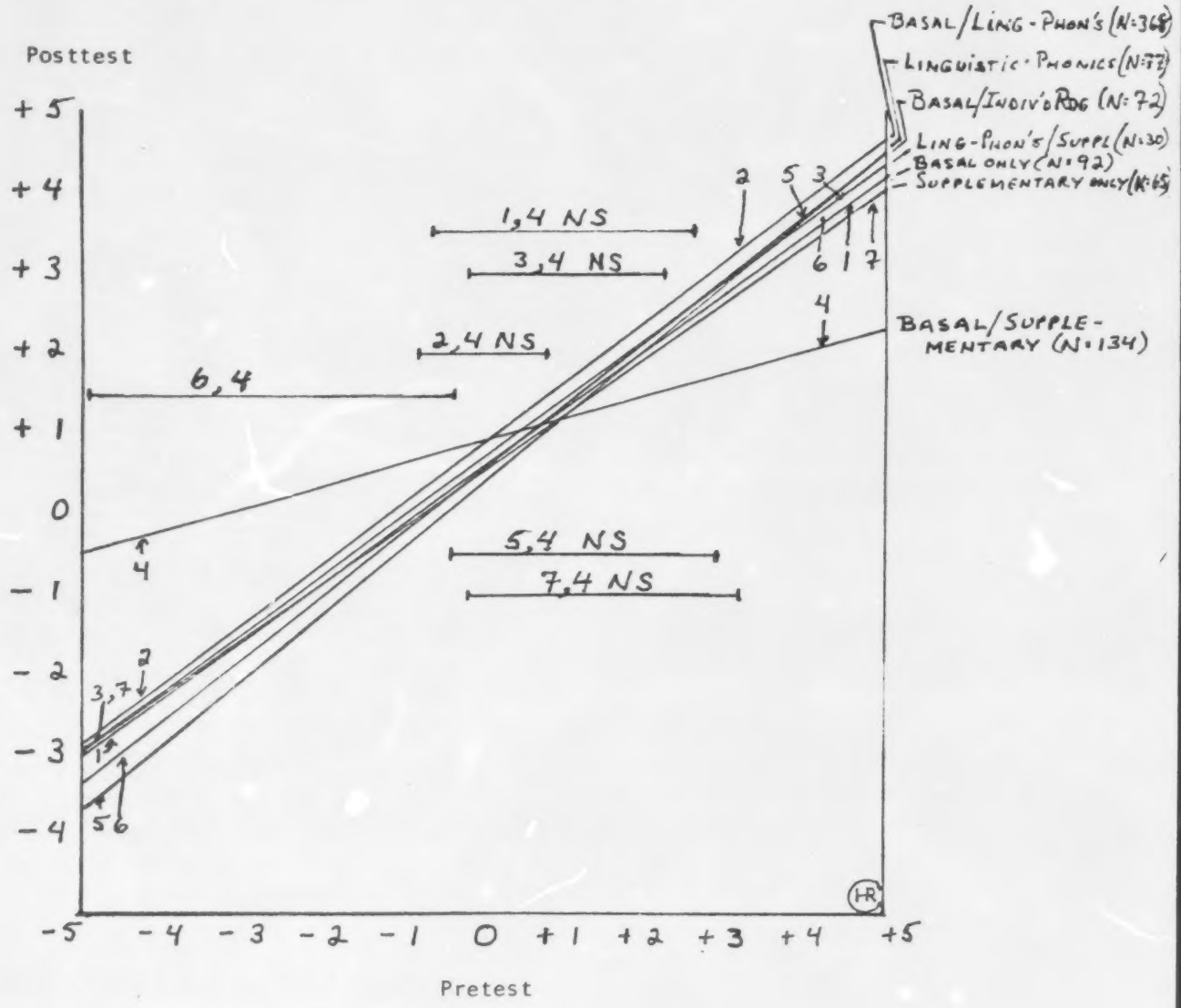
Table XXXIII

Composition of Samples for Analyses of Reading Programs, Grade 6

Programs	Number of Classes	Number of Students	Average Number of Students Per Class	Minimum Number of Students Per Class	Maximum Number Students Per Class
Basal	20	92	4.60		
Basal-Linguistic/Phonics	39	368	9.44		
Basal-Individualized Reading	17	72	4.24		
Basal-Supplementary	23	134	5.83		
Linguistic/Phonics	10	77	7.70		
Linguistic/Phonics-Supplementary	6	30	5.00		
Supplementary	5	65	13.00		
Total Number of Different Classes, Students	71	838		1	29

Figure 3

Regions of Significant Difference Between Pairs of Reading Programs, Grade 6



Note:

Because of the large number of interactions, bars are used to depict regions of significant differences or of non-significance (NS) between the regression slopes of reading programs. Numbers designate the pairs of programs to which the bars pertain.

with all the other program lines. The basic conclusion is that the Basal-Supplementary type of program was significantly more effective on the average than the other types with students who scored very low on pretest (around 35-40% objectives mastered or less, or around -.6 to -.4 logits). It was also significantly less effective on the average, compared with a pure Linguistic-Phonics type of program (2) with students who showed mastery of around 67% of the objectives or more on pretest. It is worth noting that the sample of students in the Basal-Supplementary type of program came from 23 different classes (Table XXXIII). As is evident in Figure 3, the other types of program were all comparable with each other in effectiveness, on the average.

These results are similar but not identical to those obtained in analyses of the '75-'76 data for the same types of program. In '75-'76, the Basal-Supplementary type showed the same disordinal relationship to other programs as that shown in Figure 3. On the other hand, the Basal-Linguistic/Phonics type in '75-'76 was also disordinal, closely following the Basal-Supplementary type in pattern. There were certain other regions of significance among types of programs in '75-'76 that did not show up in the '76-'77 samples. In '75-'76, the Linguistic-Phonics type was superior with good students and less effective with poor ones.

Analyses for the two years lead to the following suggestions:

- sixth grade teachers might give particular consideration to using a Basal-Supplementary type of program with students who score particularly low on pretest, but not with students who score average or above (see Evaluation System Student Reading Program Form in Appendix B for lists of Basal and of Supplementary types of materials).
- it is not yet clear how effective a Linguistic/Phonics type of program is with students who score very poorly on pretest, but it is effective with students who score average or above.
- the evidence thus far indicates that on the average the other types of program analyzed are comparable in effectiveness with students throughout the pretest range.

The results of the two-way ANCOVA were the same for both years of analysis with respect to the effect of receiving extra help in reading. There was no significant effect, nor was there a significant interaction with type of program.

1. General.

It is of interest to note that average reading level of materials increased at about the same rate, in terms of actual grade levels, at the primary grades and at the intermediate ones. The average grade levels of reading materials, by grade, are:

Grade 1: between Primer and 1^1 (grade 1, first semester)

Grade 2: slightly over 1^2 (grade 1, second semester)

Grade 3: slightly over 2^2

Grade 4: about halfway between 3^1 and 3^2

Grade 5: about halfway between 3^2 and 4

Grade 6: about halfway between 4 and 5

These, of course, are grade levels according to teachers' ratings of the materials used with the students involved in the analyses here. The average grade levels of the reading materials in the present data were slightly lower than in '75-'76 in the samples for grades 1, 2 and 6. Perhaps these shifts could have contributed to the changes in average gains noted in those grades. There is, however, not yet sufficient evidence available to make that a strong hypothesis.

Before summarizing the results of the two years of evaluations of selected types of reading programs, it is appropriate to comment on the decision criteria employed, and on the implications of the findings for subsequent analyses (e.g., analyses that will include data from school year 1977-'78).

All statistical tests have been tests of the null hypothesis, with the .05 level used as the test criterion. Any difference observed with a p-value of .050 or less was labelled significant; any differe with a p-value of greater than .050 was labelled not significant. As a general principle in this report, this has been a protected level in that if an overall test did not indicate significance, no sub-test was interpreted regardless of its p-value. The null hypothesis is often taken as the hypothesis that the observed value is probably not really different from zero, or that the observed difference could have occurred by chance 19 out of 20 times. The validity of tests of the null hypothesis rests on the assumption that samples were in fact drawn from the same population. Thus, the p-value of the statistical test is not the probability that the null hypothesis is true, given the difference observed. It is the probability of the difference occurring if the null hypothesis is true (cf. Cronbach and Snow, 1977; Carver, 1978). The probability that the null hypothesis is true, given the observed difference, is a problem for Bayesian statistics, as Cronbach and Snow point out, not for classical statistics.

In much of behavioral research, and certainly in the case of non-random samples such as those employed here, there is no way to know if the null hypothesis is true. A strict conclusion, therefore, is that the use of significance tests such as those employed here is not only irrelevant but wrong. While that conclusion is technically and theoretically correct, however, it is unsatisfactory since it gives no indication of how to proceed in decision-making in the alternative. Meanwhile, school personnel not only can but do make important decisions affecting students' lives and the utilization of resources on all manner of bases, some of which are at least as challengeable from technical and theoretical points of view as the decision process used here.

The position taken by the Evaluation System is to assume that the null hypothesis is true for samples unless there is reason to think otherwise, and

to use statistical criteria as crude pointers which essentially indicates whether or not there is a reason to take a distinction seriously. But it also insists on repeatability before suggestions or recommendations about relative effectiveness are made. Thus, the reader has seen in the preceding pages that detailed attention has been paid not just to the significance tests, but also to whether or not patterns of results have been consistent over two consecutive years. Analyses of '77-'78 data will of course give yet further evidence bearing on conclusions, but in the meantime, those consistencies that have emerged are considered sufficient to warrant making suggestions for teachers' and administrators' consideration. While the school system must do its best to avoid and ignore tenuous findings, neither can it wait until all the data are in, since all the data will never be in. Thus, the Evaluation System's approach is to try to balance theoretical purity with practicality by means of a clearly describable, conservative procedure and set of decision rules.

As a final point, the results of the '75-'76 and '76-'77 analyses of types of reading programs strongly argue for the continuation and expansion of the Evaluation System. Sufficient evidence has been found over the two years, taken together, of differential and comparable effectiveness associated with different categories or types of programs to warrant continuing with the methodology, especially as the CBC becomes widely used. For types of programs that have been found to be particularly effective or ineffective, it would now be appropriate to apply the same ANCOVA analysis to the specific reading materials employed. It is also appropriate to analyze the bases of interactions more precisely, examining variables other than just pretest. It is also appropriate to examine effects with different types of students in more detail. For example, it is appropriate now to include in analyses students who were administered different levels of the PRT pre and post. It is appropriate to

analyze the relationship of certain classroom level variables to program performance. And it is appropriate to start examining longitudinal results with the same students, as the Evaluation System is designed to do. Since there is now a basis for confidence in the ability of the Evaluation System to identify relative effectiveness and factors associated with it, there is increased justification for continuation and expansion. Expansion is important not for its own sake, but to produce samples adequate to support analyses of increasingly specifically defined programs and conditions of instruction.

J. Summary

The evaluation of types of reading programs has involved comparing the adjusted posttest reading achievement scores (percentages of reading objectives mastered) of samples of students whose teachers used the same types of reading materials. Students at each grade level were considered to have the same type of reading program if their most frequently used and next most frequently used reading materials were of the same types. These are illustrations of types of reading programs:

- Basal-Language Experience: a basal reader and a language experience material (for example: a Bank Street Reader supplemented with Experience Charts);
- Basal-Linguistic/Phonics: a basal reader and a linguistic/phonics material (for example: a Harper and Row Basic Reader supplemented by Lyons and Carnahan's Phonics We Use);
- Basal-Individualized Reading: a basal reader supplemented with an individualized reading material (for example, a Holt, Rinehart, Winston Basic Reading System material in conjunction with SRA Reading Lab material);
- Linguistic/Phonics-Supplementary: a linguistic/phonics material with a material classified as Supplementary (for example, Phono Visual supplemented with the D.C. Schools' Reading Comprehension Skills Exercises).

The Evaluation System's Student Reading Program Form in Appendix B lists a variety of materials classified under each main category or type. Table XXXIV shows by grade level the types of programs evaluated in 1975-76, in 1976-77, and in both years. For technical and professional reasons, conclusions can be drawn about differences in effectiveness, and recommendations or suggestions made, only when particular patterns of results continue for at least two consecutive years, with the pattern being statistically significant at least ones of those years. When patterns change, data from the subsequent year or years will be needed to ascertain probable trends. By those criteria, the following are the conclusions to be drawn to date from evaluations of the 1975-76/1976-77 data samples at each grade level. They pertain only to those types of reading programs with sufficient sample sizes (30 students or more) to be included in the evaluations both years.

1. Grade One

a. Four types of programs have been evaluated both in 1975-76 and in 1976-77: Basal only, Basal-Linguistic/Phonics, Basal-Programmed/Structured, and Basal-Language Experience.

b. There have been no significant differences in effectiveness on the average associated with the following three types of programs when they were compared with each other:

1) Basal only

2) Basal-Linguistic/Phonics

3) Basal-Programmed/Structured

The evidence to date indicates that teachers can use whichever type they believe is best for them and their students with the understanding that on the average these three types have been found to be comparable in effectiveness.

Table XXXIV

Types of Reading Programs Evaluated, 1975-76 and 1976-77,

by Grade*

Type of Reading Program	Grade 1	2	3	4	5	6
1. Basal only	5,6	5,6	5,6	5,6	5,6	5,6
2. Basal-Linguistic/ Phonics	5,6	5,6	5,6	5,6	5,6	5,6
3. Basal-Programmed	5,6	5,6	5			
4. Basal-Language Experience	5,6	6			5,6	
5. Basal-Individualized Reading		6	5,6	5,6	5,6	5,6
6. Basal-Supplementary	5	5,6	6	5,6	5,6	5,6
7. Linguistic/Phonics Only	6	5,6	6	6	6	5,6
8. Linguistic/Phonics Basal	5	6				
9. Ling/Phon-Indiv'd Rdg				6	5	
10. Linguistic/Phonics-Supplementary	6					6
11. Supplementary Only						6
12. Programmed-Basal		5				

*Numbers indicate year the type of program was evaluated:
5 = 1975-76
6 = 1976-77

c. The effectiveness associated with a Basal-Language Experience type of program relative to the other three listed in 1b is as yet uncertain because of a marked change in the pattern of effectiveness relationships from 1975-76 to 1976-77. Analysis of the 1977-78 data should clarify the picture.

2. Grade Two

a. The types of programs evaluated for two consecutive years were: Basal only; Basal-Linguistic/Phonics; Basal-Programmed/Structured; Basal-Supplementary; and Linguistic/Phonics only.

b. The evidence to date indicates that a pure Basal program may be more effective, on the average, than a Basal-Linguistic/Phonics, Basal-Supplementary, or a pure Linguistic/Phonics type of program.

c. The evidence to date indicates that a Basal-Linguistic/Phonics program may be more effective, on the average, than a Basal-Supplementary type of program.

d. The effectiveness relationships of a Basal-Programmed/Structured type of program to other types are uncertain at this point. Data from 1977-78 are needed to clarify them.

e. The evidence indicates that a pure Linguistic/Phonics type of program has been comparable in effectiveness, on the average, to a Basal-Linguistic/Phonics and a Basal-Supplementary type.

f. Over the two years, a program employing only a basal text, or two different basals, has consistently ranked high in average effectiveness, relative to the others evaluated. Basal-Supplementary and pure Linguistic/Phonics types have consistently ranked low in average effectiveness compared to the other types evaluated both years.

3. Grade Three

a. Three types of programs had adequate sample sizes to be included in the analyses both years: Basal only, Basal-Linguistic/Phonics; and Basal-Individualized Reading.

b. There is still some uncertainty about a choice between a pure Basal type on the one hand and a Basal-Linguistic/Phonics or a Basal-Individualized Reading type on the other. The data do suggest that between a pure Basal and a Basal-Linguistic/Phonics type of program, the Basal-Linguistic/Phonics types is likely to be the more effective of the two on the average with students who score below average on the PRT in the fall (about 50% or fewer objectives mastered).

c. A Basal-Linguistic/Phonics and a Basal-Individualized Reading type of reading program have been found to be comparable in effectiveness, on the average, over the two years.

4. Grade Four

a. Four types of program have been compared both years: Basal only, Basal-Linguistic/Phonics, Basal-Individualized Reading, and Basal-Supplementary.

b. The evidence over the two years suggests that Basal-Linguistic/Phonics and Basal-Individualized Reading types of programs are apt to be more effective, on the average, than a pure Basal program. The Basal only type of program has consistently ranked the lowest of all four types compared both years.

c. The following have been comparable in effectiveness: Basal-Linguistic/Phonics, Basal-Individualized Reading, and Basal-Supplementary.

d. The results are inconsistent as yet about the relative effectiveness associated with a pure Basal type of program compared to a Basal-Supplementary type.

5. Grade Five

a. Five types of programs have been compared both years: Basal only, Basal-Linguistic/Phonics, Basal-Language Experience, Basal-Individualized Reading, and Basal-Supplementary.

b. The findings to date indicate that of the five types, the Basal only type has consistently ranked lowest in effectiveness, on the average, for students in general. However, compared to a pure Basal program, a Basal-Language Experience type of program has given indication of being more effective with above average students (based on pretest), but less effective with below average students. That relationship also appears to be true for Basal-Language Experience in comparison to the other programs, but the 1977-78 data are needed to draw firm conclusions.

c. The following types have all been found comparable in effectiveness over the past two years of analysis (1975-76, 1976-77): Basal-Linguistic/Phonics; Basal-Individualized Reading, and Basal-Supplementary.

6. Grade Six

a. Five types of programs have been evaluated both years: Basal only, Basal-Linguistic/Phonics, Basal-Individualized Reading, Basal-Supplementary, and Linguistic/Phonics only.

b. The evidence has been consistent for two years that a Basal-Linguistic/Phonics type of program is more effective on the average than a pure Basal type, and that a pure Basal type is more effective than a Basal-Individualized Reading type of program.

c. For students who score low average or better on pretest (fall PRT), a pure Basal program has been found to be more effective than a Basal-Individualized Reading type, and a Basal-Individualized

Reading type in turn has been found to be more effective than a Basal-Supplementary type.

d. Teachers considering a choice between a Basal-Supplementary and a pure Linguistic/Phonics type of reading program should be aware that for students who have scored below average in the fall on the PRT, a pure Linguistic/Phonics type of program has been more effective on the average than a Basal-Supplementary type, while the opposite has been true for students who scored above average in the fall.

e. Effectiveness relationships are still unclear for the following program choices:

    1) Basal only versus Linguistic/Phonics only
    2) Basal-Linguistic/Phonics versus Basal-Supplementary
    3) Basal-Linguistic/Phonics versus Linguistic/Phonics only
    4) Basal-Individualized Reading versus Linguistic/Phonics only (the latter appears more effective, but whether for all students, on the average, or just for above average students can not yet be determined).

7. The results of analyses of the 1975-76 and 1976-77 data suggest that students who have received extra help in reading have gained at the same rate as those who have not. There is indication that at least at some grade levels the effectiveness of extra help may depend on the main type of reading program that the student receives. Further analyses may clarify the picture. In any case, students who have received extra help have been those who scored below those who did not, on the average, on the fall PRT. Thus, while gaps between those two categories of students have generally not closed at posttest, neither have they widened.

8. Decision-makers should bear in mind that all conclusions about program effectiveness have been based on analyses of samples of students who were given the same level of PRT in the fall and in the spring. Future evaluations will include students who were given different pre-post PRT test levels. Future evaluations will also take into account more specific characteristics of reading programs than just the general types of materials, reading level of materials, and pretest scores of students. At this stage in the development of the Evaluation System, evaluations have been limited simply to comparisons of major types of reading programs that are the most prevalent in the D.C. Schools. It is assumed that different types of programs imply different theories and methods of reading instruction. The assumption, however, has not yet been tested empirically by the Evaluation System. Finally, no distinction has been made within program samples at a given grade level between students who received different levels of the PRT. Since gain is correlated negatively with test level, different mixes of levels within programs could affect comparative performance outcomes. Future analyses will examine this potential source of bias.

9. In making decisions about choices of reading programs, decision-makers should therefore seriously consider the findings to date but not feel compelled to act upon them mechanically in all cases. As noted throughout, the findings are based on average outcomes. Furthermore, the decision-maker needs to keep in mind that statements about the effectiveness of types of reading programs mean average effectiveness associated with them, not necessarily caused by them. Other data considered pertinent and important in making program choices should also be taken into consideration.

VII. CONCLUSIONS

Data from the 34 elementary schools in the Evaluation System were analyzed separately for the school years 1975-76 and 1976-77. As with the 1975-76 report, the 1976-77 findings have been presented in a two-part report: The summary report and the technical report. From these findings main conclusions were drawn. Where appropriate, the findings from the 1975-76 report have been incorporated in this report.

A. Reading Achievement and Testing

1. In 1976-77 the average percentage of total possible reading objectives mastered in the fall on the Prescriptive Reading Test (PRT) by 376 classes of students in grades 1-6 was 59%. The average percent of total possible reading objectives mastered by the same students the following spring was 74%. Thus, an additional 15% of the total number of reading objectives on the PRT were mastered by the average elementary class during the 1976-77 school year.
2. The highest average gain in percentage of reading objectives mastered was found in second grade classes (21%), the lowest in fifth grade classes (9%). Generally, average gains were higher for primary grades than for intermediate grades.
3. There were no significant differences in average percentage of objectives mastered on the pretests for 1975-76 and 1976-77.
4. There was a statistically significant difference in the average percentage of objectives mastered at posttest by the 1976-77 sixth grade classes over the 1975-76 sixth grade classes. The sixth grade average gain in 1975-76 was 5.4%;

it was 10.3% in 1976-77. Average posttest gains for the other grades revealed no significant differences.

5. Across all six grades, results show that the greatest pre-post gain in reading scores is seen in classes which:

   a) have higher average interest in reading;

   b) use lower levels of reading materials;

   c) meet in self-contained classrooms, as opposed to open classrooms. (This last result was not present in 1975-76 data and should be studied further before action is taken.)

6. For the primary grades alone, significantly greater pre-post reading score gains were seen in classes having students who do a larger amount of outside reading.

7. For the 376 classes studied at all grade levels, teachers administered an average of 2.0 different levels of the PRT per class in the fall and spring testings. That is a minimum average, since it is based on a count only of students who were given the same level of test at both testings.

B. Interest in Reading

As was noted above, higher interest in reading, on a class average basis, contributes to higher achievement in reading. It therefore becomes particularly important to address those factors which tend to increase reading interest. Results show that class average interest in reading is significantly greater in classes which:

a. have lower mean absence rates, particularly in the intermediate grades;

b. have lower percentages of male students, particularly in the primary grades;

c. have low student turnover rates, particularly in the primary grades;

d. have a low need for extra help in reading;

e. receive lesser amounts of extra help in reading;

f. have smaller numbers of reading problems;

g. use team teaching approaches particularly in the primary grades;

h. use lower levels of reading material.

C. Evaluations of Selected Types of Reading Programs

1. Commonly used types of reading program at each grade have been compared for two consecutive years (1975-76 and 1976-77). Using the results of both analyses it has been possible to suggest guidelines to be considered by decision-makers in selecting types of reading programs. The guidelines suggested for each grade level are listed under "Findings" in Section VI of this Summary Report.

2. The guidelines, which are based on the application of a proposed set of decision rules for interpreting the results of evaluations for two consecutive years, consist of three kinds of statements:

a. an identification of types of reading programs which have consistently appeared to be more, or less, effective than other types for specified types of students;

b. an identification of types of reading programs which have consistently appeared to be comparable in effectiveness for specified types of students;

c. an identification of types of programs about which the evidence is still unclear because of marked changes in relative effectiveness from one year to the next (further evaluations are needed in subsequent years to clarify the effectiveness relationships of these types of programs).

3. All guidelines have come from the analysis of pre and posttest PRT scores of students who were administered the same level of test in the fall and in the spring. They have come from measures of effectiveness consistently associated with different types of reading programs but not necessarily caused by them. Future evaluations of effectiveness will take into consideration students who were administered different levels of the PRT in the fall and in the spring. They will also take into consideration additional factors, other than pretest score and the grade level of the reading materials used, that may influence the observed comparative effectiveness of different types of reading programs with different types of students.
4. The present guidelines should be considered by decision-makers along with other pertinent and important information when they are making decisions about the types of reading programs to use at each grade level. The guidelines are not to be viewed as rigid prescriptions, but as aids to decision-making.
5. The methods of the Evaluation System have been found to yield results that make the continuation and expansion of the Evaluation System warranted. Continuation and expansion can be expected to lead to more detailed and specific guidelines for choosing and implementing reading programs.

[redacted] SCHOOL [redacted] E:

[redacted] UNIT.[redacted]

TEACHING SITUATION

(a) (c) (e) (g) (i)
(b) (d) (f) (h) (j)

CLASSROOM

- Regular
- TAIR
- Other

IS AIDE

- Paid?
- Volunteer?

ROOM

GRADE

- Pre K
- K
- 1
- 2
- 3
- 4
- 5
- 6
- 7
- 8
- 9
- 10
- 11
- 12
- Ungraded K-6
- Ungraded K-3
- Ungraded 4-6
- 3-7 year olds
- 4-8 year olds
- 8-10 year olds
- 10-12 year olds

AIDE HOURS PER WEEK

TEACHER'S NAME (Last, First, Initial)

AIDE'S NAME (Last, First, Initial)

[redacted] a [redacted] g re-source teacher available?

- Yes
- No

Is there a math resource teacher available?

- Yes
- No

NUMBER FREE LUNCH STUDENTS

 # OF APPLICATIONS APPROVED FOR FREE LUNCH

 # OF REDUCED COST LUNCH APPLICATIONS APPROVED

TOTAL MEMBERSHIP IN CLASS

SCHOOL NO.

TEACHER NUMBER

START

END DAY

ABSENT

Sem. 1	Sem. 2

AIDE NUMBER

START

END DAY

ABSENT

Sem. 1	Sem. 2

[redacted]

STUDENT NAME

ENTRY

E1 E4b
E2 E5
E3 E6
E4a Unk

EXIT

D7a D10
D7b D11a
D8 D11b
D9 Unk

SEX

Male
Female

BIRTH DATE

Month Year

STUDENT NUMBER

START DAY

END DAY

ABSENT

Sem. 1 Sem. 2

SCHOOL

STUDENT NAME

ENTRY

E1 E4b
E2 E5
E3 E6
E4a Unk

EXIT

D7a D10
D7b D11a
D8 D11b
D9 Unk

SEX

Male
Female

BIRTH DATE

Month Year

STUDENT NUMBER

START DAY

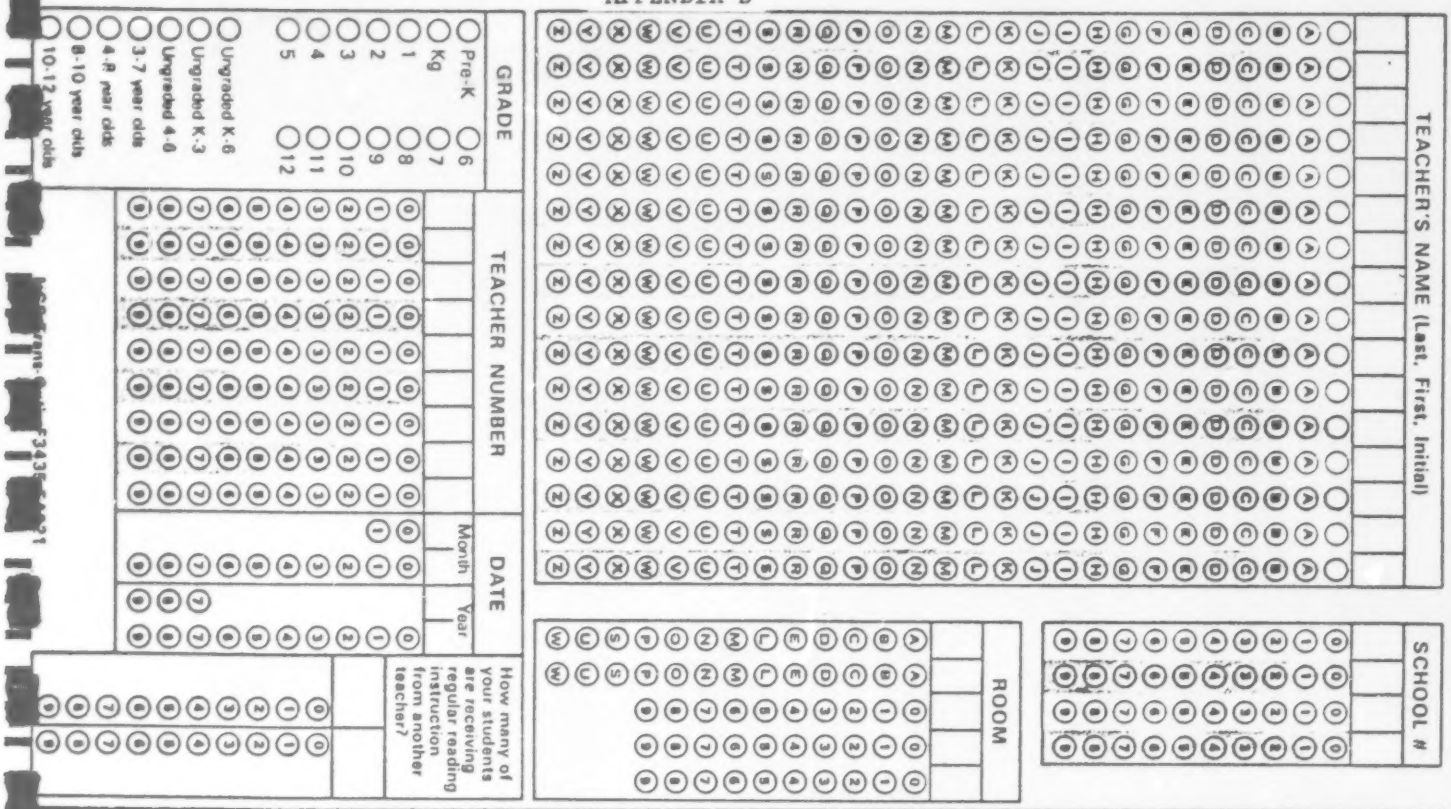
END DAY

ABSENT

Sem. 1 Sem. 2

SCHOOL

APPENDIX B



The following questions pertain to your family or home-room class.

Often students have needs or problems that affect the teacher's instructional program or the students' ability to learn:

a. As a result, have you referred any of your students this year to any of the following?

- ○ Principal
- ○ Counselor
- ○ Health Specialist (Nurse)
- ○ Speech Therapist
- ○ Pupil Personnel Worker
- ○ Visiting Instructor
- ○ Special Teacher(s) (Reading, Math Specialists, etc.)
- ○ School Psychologist
- ○ Social Worker
- ○ Aides (Educational, Title I, Health, Community, etc.)
- ○ None
- ○ Other (Specify) _____

b. Please mark the number of students you referred this year because they had learning problems in one of the following categories: Reading; Math; Speech; Language Arts; Other

	1-3	4-6	7-9	10 or more
Reading	○	○	○	○
Math	○	○	○	○
Speech	○	○	○	○
Language Arts	○	○	○	○
Other	○	○	○	○

136

c. Please mark the number of students you referred this year because they had learning barriers in one of the following categories:

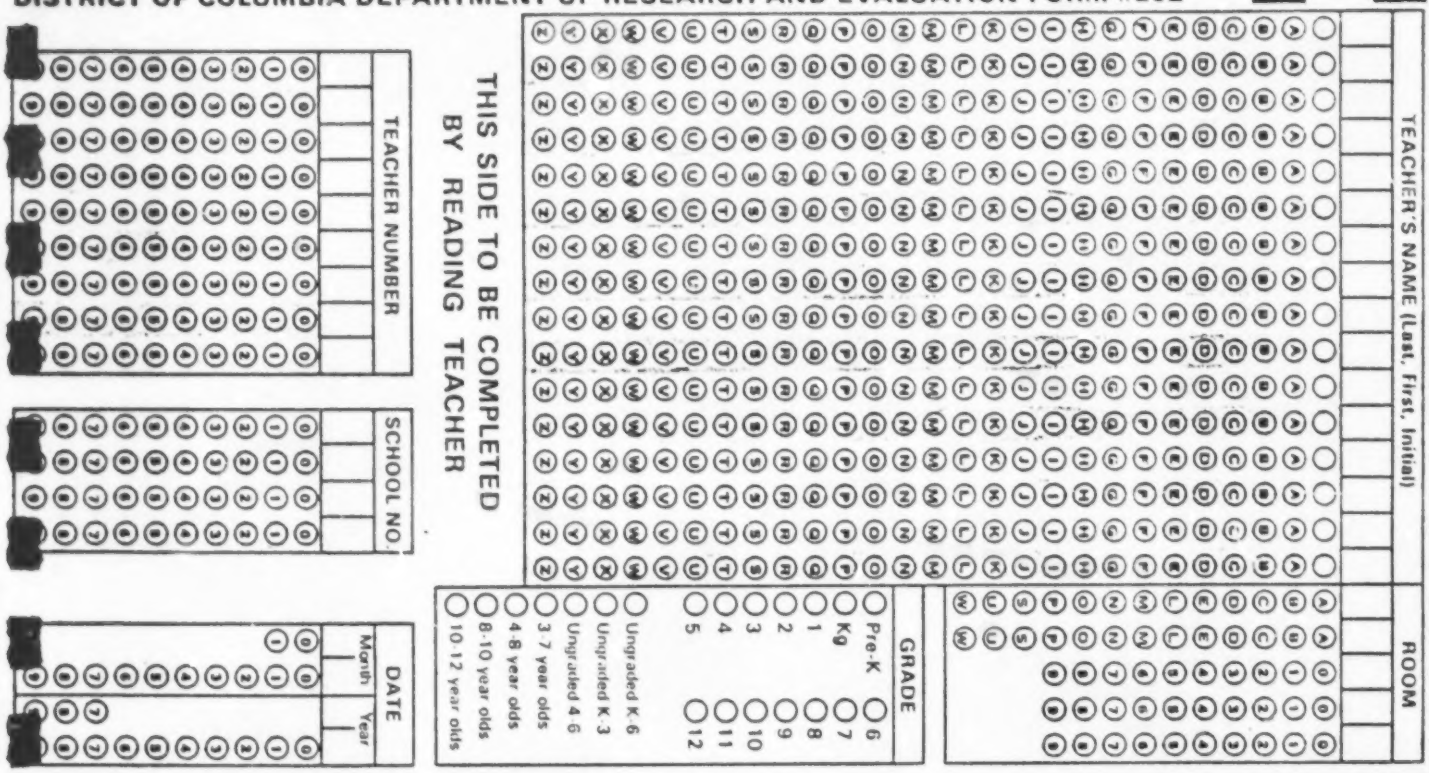
	1-3	4-6	7-9	10 or more
Behavior*	○	○	○	○
Attendance	○	○	○	○
Truancy	○	○	○	○
Nutrition	○	○	○	○
Drugs	○	○	○	○
Home or Community	○	○	○	○
Child Abuse	○	○	○	○
Economic Conditions (Clothing, shoes, etc.)	○	○	○	○
Medical/Physical Health (Vision, Sickness, Dental, Pregnancy, Injury, etc.)	○	○	○	○

d. (Answer only if you have made NO referrals)
Why haven't you made referrals?

Learning Problems
Learning Barriers

- ○○ None of my students had a need to be referred
- ○○ I prefer to deal with the problems of my students myself if at all possible.
- ○○ The necessary special teachers were not available
- ○○ There was no counselor available
- ○○ There was no health specialist available (nurse, health aide, etc.)
- ○○ There is a lack of response to referrals
- ○○ I lack confidence in our referral system
- ○○ I lack confidence in the resource personnel.
- ○○ The referral procedures are poor.
- ○○ We are encouraged by the school's policy to handle our own classroom problems
- ○○ Other

DISTRICT OF COLUMBIA DEPARTMENT OF RESEARCH AND EVALUATION FORM #202



1. How many regular reading classes do you teach during the school day?
   - ○ one
   - ○ two
   - ○ three
   - ○ four
   - ○ five
   - ○ six
   - ○ seven
   - ○ eight

2. If you teach only one reading class during the day, during what time of the day is it taught?
   - ○ before 10 am
   - ○ between 10 am and noon
   - ○ afternoon

3. How long have you been the reading teacher for your class or classes this year?
   - ○ less than 1 month
   - ○ 1 month
   - ○ 2-3 months
   - ○ 4-5 months
   - ○ 6-7 months
   - ○ 8-9 months

4. Have you been teaching your reading class or classes by a team teaching approach this year?
   - ○ yes
   - ○ no
   - ○ no, but we exchange students for reading

5. Have you taught any reading skills objectives for your class this year other than those on the attached list?
   - ○ yes
   - ○ no

6. If yes, what percent of the reading skills objectives you taught were not contained in the attached PRT listing?
   - ○ less than 10%
   - ○ 10% to 20%
   - ○ 21% to 30%
   - ○ 31% to 40%
   - ○ 41% to 50%
   - ○ 51% to 60%
   - ○ 61% to 70%
   - ○ 71% to 80%
   - ○ over 80%

7. Did you teach PRT skills objectives from more than one of the test levels this year?
   - ○ yes
   - ○ no

8. Were you free to choose the reading skills objectives taught in your class(es) this year?
   - ○ totally
   - ○ to a great extent
   - ○ to some extent
   - ○ no

9. To what extent do you coordinate your instructional reading program with others in the school who also work with any student in your reading class or classes? (Mark all those that apply)

OTHER PERSONS WHO HELP MY STUDENTS IN READING

totally	to a great extent	to some extent	not at all	
○	○	○	○	The child's family or homeroom teacher (if you are not that person)
○	○	○	○	The reading specialist(s)
○	○	○	○	The reading resource teacher(s)
○	○	○	○	The reading aide(s)
○	○	○	○	The reading tutor(s)
○	○	○	○	Others (Specify) _____

130

Dist. of Col. Dept. of Research and Evaluation #200

[redacted]

Teacher Name _____ School Name _____

STUDENT NAME

STUDENT ID NUMBER

TEACHER NUMBER

SCHOOL

Teacher

Aide

Section

If you are not this student's family or homeroom teacher, please print your name and the name of the aide who helps you in the reading class, and group or section of this student if you have more than one group.

B. 1. How many months has the child been in your class this year?

○ Less than one month ○ 1 month ○ 2 months ○ 3 months ○ 4 months
○ 5 months ○ 6 months ○ 7 months ○ 8 months

B. 2. Is this child in need of extra help in reading?

○ Yes, very much ○ Yes, somewhat ○ No

B. 3. THIS YEAR from what sources OTHER THAN YOU has this child received reading instruction DURING THE REGULAR SCHOOL DAY? Please mark all applicable.

○ None, no extra help ○ College student tutor ○ Other elem. sch. student tutor ○ Parent tutor ○ Don't know - new student
○ Reading specialist/reading resource teacher ○ Teacher aide tutor ○ Jr. high student tutor ○ Urban service corps tutor
○ Service group tutor ○ Sr. high student tutor ○ Other

B. 4. To the best of your knowledge how many months has this help been provided by the sources marked above?

○ Less than one month ○ 1 month ○ 2 months ○ 3 months ○ 4 months
○ 5 months ○ 6 months ○ 7 months ○ 8 months

B. 5. How often each week does this child receive reading instruction from the sources listed above? Mark most appropriate.

○ 5 Days ○ 4 Days ○ 3 Days ○ 2 Days ○ 1 Day ○ Less than once a week

B. 6. If this child needs extra help in reading and is not receiving it, please mark the main reason.

○ No reading specialist/reading resource teacher ○ Lack of instructional materials ○ Other (please explain) _____
○ Lack of tutorial services ○ Lack of teacher aides _____
○ Lack of diagnostic information

D. Why are you using the materials listed on the reverse side of this form for this student?

○ They are the materials I have found most effective in the past.
○ They are the only materials most readily available in my school.
○ They are materials suggested or recommended by the principal or supervisor.
○ They are keyed to the school system reading objectives and criterion-referenced tests.
○ Other
DO NOT WRITE OUTSIDE BOX
○ They are materials found most suitable to this student's needs.
○ I learned about the materials in an in-service program/workshop.
○ I selected them on the basis of my own diagnostic test and observations.
○ I selected them on the basis of standardized test results.
○ They are the materials used as part of the special program in this building.

C. Aside from reading instruction, how often does this student read on his own in school? (library books, comics, etc.)

○ Often ○ Sometimes ○ Never

E. How would you assess this child's interest in reading?

○ Very interested
○ Moderately interested
○ Seldom interested
○ Not interested
○ Actively dislikes reading

100

NCS Trans-Optic F3369-54321

EVALUATION SYSTEM
STUDENT READING PROGRAM FORM

MARKING INSTRUCTIONS

1. Use black lead pencil only (No. 2½ or softer.)
2. Do not use ink or ballpoint pens.
3. Make heavy black marks that fill the circle completely.
4. Erase completely any mark you wish to change.
5. Do not make any stray marks on this sheet.

BASAL SERIES

101 Bank Street Readers – MacMillan
102 Basic Curriculum Series – Lippincott (McLeod, McCracken & Walcutt)
103 Betts Basic Readers – American Book Co.
104 Bookmark Reading Program – Harcourt, World & Brace
105 Ginn Basic Readers, 100 Edition (Russell, et. al.)
106 Ginn Basic Readers, 360 Program (Clymer, et. al.)
107 Harper and Row Basic Readers (Van Roekel, et. al.)
108 Houghton Mifflin – Reading for Meaning (McKee)
110 Lyons and Carnahan – Curriculum Motivation Series (Fay)
111 MacMillan Reading Program (Harris & Clark)
112 Reading Caravans – D. C. Heath Co. (Witty and Freeland)
113 Reading Experience and Development (Read) – American Book Co.
114 Scott, Foresman – Bright Horizons Reading Program
115 Scott, Foresman – Curriculum Foundation Series
116 Scott, Foresman – Open Highways Program
117 Sheldon Basic Readers – Allyn & Bacon
118 Cambridge Book Co. – Reading Series
119 D. C. Heath – Communications
120 Ginn Basic Readers, 300 Edition
121 Holt, Rhinehart, Winston – Basic Reading System
122 Houghton Mifflin – Interaction
123 MacMillan – Series R
124 New Dimensions in Education – Alfa Series
125 Scott, Foresman – Reading Systems
199 Other (Please specify when gridding in box below)

LINGUISTIC/PHONIC

201 Basic Reading Series – SRA – (Pig in a Jig Books)
202 Lippincott – Basic Reading
203 Merrill Linguistic Readers (Fries, Wilson & Rudolph)
205 Open Court Foundation Program
206 Palo Alto Reading Program – Harcourt, Brace & World
207 Phonetic Keys to Reading Series – The Economy Co.
208 Phonics We Use – Lyons and Carnahan
209 Phono Visual – Phono Visual Products
211 Readiness in the Language Arts (RLA) – Cynthia Buchanan
299 Other (Please specify when gridding in box below)

PROGRAMMED/STRUCTURED

301 Lift-Off to Reading – SRA – (Woolman)
302 Programmed Reading – Webster Division of McGraw Hill – (Buchanan & Sullivan)
303 Project READ – Behavioral Research Labs – (Sullivan's Reading)
304 Starter/101 – (Silver Burdett (O'Keefe)
305 Random House – Structural Reading
306 Reading Readiness Tapes – Ideal School Supply
399 Other (Please specify when gridding in box below)

LANGUAGE EXPERIENCE

401 Experience Charts – Commercial
402 Experience Charts – Made by Teacher with Class
403 Language Experience Readers – Noble & Noble (Chandler Reading Program)
499 Other (Please specify when gridding in box below)

INDIVIDUALIZED READING

501 Extensive Use of Variety of Published Materials
502 Little Owl Series (Classroom Library) – Holt, Rinehart, Winston
503 MacMillan Reading Spectrum
505 SRA Reading Labs
506 Extensive Use of Variety of Non-Published Materials
599 Other (Please specify when gridding in box below)

SUPPLEMENTARY MATERIALS

605 Invitations to Personal Reading (Classroom Library) – Scott, Foresman
609 Reading Comprehension Skills Exercises (D. C. Schools)
610 Readers Digest Services, Inc.
611 Singer, Random House Literature Series
612 Torch Lighter Kits
613 Weekly Readers
614 Teacher Made Materials
615 Other (Please specify when gridding in box below)

SPECIAL

702 Distar – SRA (Englemann)
703 Frostig's Developmental Program in Visual Perception – Follett
704 I/T/A
705 Kindergarten Keys – Economy Co.
799 Other (Please specify when gridding in box below)

MOST USED MATERIAL	2nd MOST USED MATERIAL	3rd MOST USED MATERIAL
Approx. Level	Approx. Level	Approx. Level
○ RR ○ 3²	○ RR ○ 3²	○ RR ○ 3²
○ PP ○ 4	○ PP ○ 4	○ PP ○ 4
○ P ○ 5	○ P ○ 5	○ P ○ 5
○ 1¹ ○ 6	○ 1¹ ○ 6	○ 1¹ ○ 6
○ 1² ○ 7	○ 1² ○ 7	○ 1² ○ 7
○ 2¹ ○ 8	○ 2¹ ○ 8	○ 2¹ ○ 8
○ 2² ○ 9	○ 2² ○ 9	○ 2² ○ 9
○ 3¹ or above	○ 3¹ or above	○ 3¹ or above
OTHER (199, 299, ETC.) SPECIFY	OTHER (199, 299, ETC.) SPECIFY	OTHER (199, 299, ETC.) SPECIFY

A. To what extent is the reading program described above indicative of what was used all year for this child for reading instruction? (Exclude changes in reading level)

○ Don't know - student new to this class
○ Very representative or very similar
○ Somewhat representative or somewhat similar
○ Not at all representative or not at all similar

19_

APPENDIX C

APPROVED LIST OF PARTICIPANT SCHOOLS
IN THE SCHOOL YEAR 1975-76 EVALUATION SYSTEM

Region I (N=7)	Beers
	Congress Heights
	Friendship
	Hendley
	McGogney
	Malcolm X
	Savoy
Region II (N=9)	Amidon
	Bowen
	Hardy
	Hyde
	Key
	Lafayette
	Mann
	Stoddert
	Syphax
Region III (N=4)	Blow (Now Region IV)
	Davis
	Miner (Now Region IV)
	Thomas
Region IV (N=5)	Benning (Now Region III)
	Goding
	Walker-Jones
	Weatherless (Now Region III)
	Wilson, J.O.
Region V (N=6)	Bancroft (Now Region II)
	Bruce-Monroe
	Bunker Hill
	Cleveland
	Meyer
	Petworth
Region VI (N=3)	Noyes
	Shaed
	Whittier

Total Schools = 34

APPENDIX D

Table A

Correlation Matrix Across All-Grades Regression on Posttest Scores

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
(1) Grade	1.00000	.18974	.07604	-.21916	.07681	-.22347	.11347	-.12870	.00266	.07656	.32065	-.12474
(2) Test Level Count	.18974	1.00000	.04830	-.03617	.04404	-.07653	.20214	-.02240	.05570	.28125	.28206	.30444
(3) Pretest Score	.07604	.04830	1.00000	.68273	.99936	.67317	.00973	-.16636	-.10122	.07128	.15735	-.11195
(4) Raw Pretest	-.21916	-.03617	.68273	1.00000	.68244	.99224	.01340	-.12655	-.08422	.06164	-.14874	-.15017
(5) Arc SQRT Pretest	.07681	.04404	.99936	.68244	1.00000	.67537	.00844	-.15777	-.10280	.06859	.15091	-.11185
(6) Raw Posttest	-.22347	-.07653	.67317	.99224	.67537	1.00000	.00457	-.12231	-.08028	.03944	-.15091	-.16028
(7) Arc SQRT Posttest	.11347	.20214	.00973	.01340	.00844	.00457	1.00000	-.07112	.02803	.25018	.06155	.05177
(8) Class Size	-.12870	-.02240	-.16636	-.12655	-.16777	-.12231	-.07112	1.00000	.07986	-.08227	-.18157	.14670
(9) Class Absence	.00266	.05570	-.10122	-.08422	-.10280	-.08028	.02303	.07986	1.00000	.02605	-.05132	.04024
(10) % Males	.07656	.28125	.07128	.06164	.06859	.03944	.25018	-.08227	.02605	1.00000	.16663	-.03693
(11) Class Turnover	.32065	.28206	.15735	-.14874	.16031	-.15091	.06155	-.18157	-.05132	.16663	1.00000	-.13273
(12) Mean Material Level	-.12474	.30444	-.11195	-.15017	-.11185	-.16028	.05177	.14670	.04024	-.03693	-.13273	1.00000
(13) Variability Material Level	.10784	.08770	-.21936	-.32778	-.22016	-.34196	-.01542	.19803	.10549	-.02446	.00531	.16309
(14) Reading on Own	.03483	.07189	-.23531	-.32901	-.23705	-.33940	.01274	.17495	.04538	-.10784	-.06456	.22763
(15) Reading Interest	-.12382	-.07989	.30539	.30305	.30619	.30565	-.00673	-.26171	-.12879	.09385	.27942	-.23751
(16) Need for Help	.13103	.04538	.12931	.03833	.13258	.04262	.00199	-.09894	-.08096	.05165	.15768	.00676
(17) Teacher Turnover	.00747	-.05971	.05404	-.00715	.05394	-.01208	-.07414	.04186	.04682	-.04789	-.01053	.02574
(18) Teacher Absence	.04970	.04430	-.00102	-.05942	-.00070	-.05808	.03236	-.06493	-.02648	-.06238	.00692	.03703
(19) % Free Lunch	-.04085	-.00962	-.32140	-.30568	-.32011	-.29362	-.19270	.20070	-.03221	-.21475	-.16001	.09669
(20) Math Specialist	.14556	-.00412	.17488	.12923	.17389	.11540	-.06277	-.14232	-.00621	.14856	.15344	-.15188
(21) Team Teaching	-.23307	.12569	-.22911	-.22265	-.22931	-.22511	.06230	.12388	.04616	-.03767	-.16871	.28192
(22) Open Space	-.16982	.07674	-.10515	-.13481	-.10720	-.14167	.02905	.13696	-.02661	-.08369	-.09261	.23394
(23) Reading Problems	-.06819	-.06597	-.10686	-.05907	-.10601	-.05361	-.05754	.22707	.01935	-.04679	-.16057	.10271
(24) Behavior Barriers	.01643	-.02409	-.12174	-.17509	-.12023	-.16925	-.00567	.06842	.09621	-.00175	-.01572	.12240
(25) Health Barriers	-.13222	-.02449	-.01380	.03187	-.01374	.03266	-.07566	.05605	.05992	.01746	-.12620	.10915

	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)	(22)	(23)	(24)
(1) Grade	.10784	.03483	-.12382	.13103	.00747	.04970	-.04085	.14556	-.23307	-.16982	-.06819	.01643
(2) Test Level Count	.08770	.07189	-.07989	.04538	-.05971	.04430	-.00962	-.00412	.12569	.07674	-.06597	.02409
(3) Pretest Score	-.21936	-.23531	.30539	.12931	.05404	-.00102	-.32140	.17486	-.22911	-.10515	-.10686	-.12174
(4) Raw Pretest	-.32778	-.32901	.30305	.03833	-.00715	-.05942	-.30568	.12423	-.22265	-.13481	-.05907	-.17509
(5) Arc SQRT Pretest	-.22016	-.23705	.30619	.13258	.05394	-.00070	-.32011	.17389	-.22931	-.10720	-.10601	-.12023
(6) Raw Posttest	-.34196	-.33940	.30565	.04262	-.01208	-.05808	-.29362	.11540	-.22511	-.14167	-.05361	-.15925
(7) Arc SQRT Posttest	-.01542	.01274	-.00673	.00199	-.07414	.03236	-.19270	-.06277	.06230	.02905	-.05754	.00567
(8) Class Size	.19803	.17495	-.26171	-.08894	.04186	-.06493	.20070	-.14232	.12388	.13696	.22707	.05842
(9) Class Absence	.10849	.17841	-.12879	-.08096	.04682	-.02648	-.03221	-.00621	.04616	-.02661	.01935	.09821
(10) % Males	-.02446	-.10784	.09385	.05165	-.04789	.06238	-.21475	.14856	-.03767	-.08369	-.04679	-.00175
(11) Class Turnover	.00531	-.06456	.27942	.15768	-.01053	.00692	-.16001	.15344	-.16871	-.09261	-.16057	-.01572
(12) Mean Material Level	.14309	.22763	-.23751	.00676	.02574	.03703	.09669	-.15138	.28192	.23394	.10271	.12240
(13) Variability Material Level	1.00000	.75552	-.40083	.04747	-.06171	.08528	.24897	-.09308	.19473	.20799	.02389	.10875
(14) Reading on Own	.75552	1.00000	-.39659	.04450	-.03843	.04247	.17606	-.14267	.10690	.15539	.00334	.14375
(15) Reading Interest	-.40083	-.39659	1.00000	.12382	.06380	-.03988	-.29324	.17129	-.20659	-.14037	-.27499	-.20291
(16) Need for Help	.04747	.04450	.12382	1.00000	.01107	.07896	-.08031	.04953	-.12250	-.17420	-.18258	.03448
(17) Teacher Turnover	-.06171	-.03843	.06380	.01107	1.00000	.01079	.01984	-.03402	-.10344	-.08528	-.00719	-.08238
(18) Teacher Absence	.08528	.04247	-.03988	.07896	.01079	1.00000	.05985	.01520	.03435	-.02252	-.03256	.00277
(19) % Free Lunch	.24897	.17606	-.29324	-.08031	.01984	.05985	1.00000	-.18584	.13335	.06220	.21033	.04193
(20) Math Specialist	-.09308	-.14267	.17129	.04953	-.03402	.01520	-.18584	1.00000	-.29769	-.34499	-.18601	-.05162
(21) Team Teaching	.19473	.10890	-.20859	-.12250	-.10344	.03435	.13335	-.29769	1.00000	.55512	-.00413	.07074
(22) Open Space	.20799	.15539	-.14037	-.17420	-.08528	-.02252	.06220	-.34499	.55512	1.00000	.05651	-.04134
(23) Reading Problems	.02389	.00334	-.27499	-.18258	-.00719	-.03256	.21033	-.18601	-.00413	.05651	1.00000	.20182
(24) Behavior Barriers	.10875	.14375	-.20291	.03448	-.08238	.00277	.04193	-.05162	.07074	-.04134	.20182	1.00000
(25) Health Barriers	-.01381	-.02211	.17530	-.10215	-.06747	.02933	.08611	-.07424	.09452	.05374	.26540	.24352

		(25)
(1)	Grade	-.13222
(2)	Test Level Count	-.02449
(3)	Pretest Score	-.01380
(4)	Raw Pretest	.03197
(5)	Arc SQRT Pretest	-.01374
(6)	Raw Posttest	.03266
(7)	Arc SQRT Posttest	-.07566
(8)	Class Size	.05605
(9)	Class Absence	.05992
(10)	% Males	.01746
(11)	Class Turnover	-.12620
(12)	Mean Material Level	.10915
(13)	Variability Material Level	-.01381
(14)	Reading on Own	-.02211
(15)	Reading Interest	-.17530
(16)	Need for Help	-.10215
(17)	Teacher Turnover	-.06747
(18)	Teacher Absence	.02933
(19)	% Free Lunch	.08611
(20)	Math Specialist	-.07424
(21)	Team Teaching	.09452
(22)	Open Space	.05374
(23)	Reading Problems	.26540
(24)	Behavior Barriers	.24352
(25)	Health Barriers	1.00000

APPENDIX E

Table B

Correlation Matrix Primary Grades

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
(1) Grade	1.00000	-.28154	.26885	-.28728	.27332	-.29420	.00546	-.35510	-.04036	-.15979	.45235	-.17141
(2) Test Level Count	-.28154	1.00000	-.06612	-.13113	-.06907	-.15043	.09427	.03503	.09416	.23422	.24820	-.46013
(3) Pretest Score	.26885	-.06612	1.00000	-.68146	.99919	.68071	.05094	-.19926	-.18966	-.00137	.30958	-.03273
(4) Raw Pretest	-.28728	-.13113	.68146	1.00000	-.68154	.99023	.13239	-.11806	-.12135	-.03328	.19138	-.15757
(5) Arc SQRT Pretest	.27332	-.06907	.99919	-.68154	1.00000	-.68425	-.04828	-.20146	-.19176	-.00302	.31717	-.03554
(6) Raw Posttest	-.29420	-.15043	.68071	.99023	-.68425	1.00000	.11185	-.11351	-.11864	.01387	.20486	-.17783
(7) Arc SQRT Posttest	.00546	.09427	.05094	.13239	-.04828	.11185	1.00000	-.00693	.13626	.32305	-.02355	-.00437
(8) Class Size	-.35510	.03503	-.19926	-.11806	-.20146	-.11351	-.00693	1.00000	.05944	.02036	-.16702	.05993
(9) Class Absence	-.04036	.09416	-.18966	-.12135	-.19176	-.11864	.13626	.05944	1.00000	.04258	-.07523	.08417
(10) % Males	-.15979	.23422	-.00137	.03328	-.00302	.01387	.32305	.02036	.04258	1.00000	.10116	.07050
(11) Class Turnover	.45235	.24820	.30958	.19138	.31717	.20486	-.02355	-.16702	-.07523	.10116	1.00000	.14378
(12) Mean Material Level	-.17141	.46013	-.03273	-.15757	-.03554	-.17783	-.00437	.05993	.08417	.07050	.14378	1.00000
(13) Variability Material Level	-.12545	.02389	-.32166	-.43626	-.32329	-.44590	-.01640	.15197	.14940	-.02160	-.09412	.10110
(14) Reading on Own	-.18757	-.00174	-.27187	-.39061	-.27520	-.40234	.02987	.11790	.24163	-.11337	-.18086	.16418
(15) Reading Interest	.22731	-.13484	.40919	.40792	.40862	.40834	-.09399	-.23900	-.24148	.00877	.30934	-.19415
(16) Need for Help	-.02748	-.06588	.14872	.06181	.15479	.07635	-.03425	-.03255	-.04221	-.02446	.01656	.03190
(17) Teacher Turnover	.08931	-.02831	.04088	-.03871	.03870	-.04826	-.02212	-.04562	-.02208	-.02446	.01656	.03190
(18) Teacher Absence	.01759	-.08462	-.03055	-.05871	-.03191	-.06422	-.04561	-.01200	.08999	-.08789	-.08789	.04253
(19) % Free Lunch	-.06718	.01452	-.38354	-.37349	-.37881	-.34838	-.18734	-.11463	-.09417	-.02573	-.16343	-.16797	-.01423
(20) Math Specialist	.10671	-.13186	.16267	.22527	.16047	.21541	-.08247	-.12357	-.03924	.08756	.10281	-.09438
(21) Team Teaching	-.40844	.35740	-.24735	-.33961	-.24895	-.34416	-.01074	.15675	.09417	-.02573	-.10793	.37806
(22) Open Space	-.32552	.27562	-.21041	-.26883	-.21382	-.27983	-.02705	.19550	.03927	-.01655	-.02422	.30866
(23) Reading Problems	.11006	-.07904	-.09978	-.06090	-.09874	-.05432	-.01200	.10128	.03194	.02525	-.07026	.04373
(24) Behavior Barriers	-.04917	.02158	-.12788	-.19646	-.12492	-.18132	-.08043	-.00285	.11322	.01748	.07450	.09800
(25) Health Barriers	.04642	.06463	-.08398	-.01536	-.08393	-.01341	-.06925	.02979	.07735	.04306	.08766	.07690

	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)	(22)	(23)	(24)
(1) Grade	-.12545	-.18757	.22731	-.02748	.08931	.01759	-.06718	.10671	-.40844	-.32852	.11006	-.04917
(2) Test Level Count	.02389	-.00174	-.13484	-.06588	-.02831	-.08462	.01452	-.13186	.35740	.27562	-.07904	.02158
(3) Pretest Score	-.32166	-.27187	.40919	.14872	.04088	-.03055	-.38354	.16267	-.24735	-.21041	-.09978	-.12788
(4) Raw Pretest	-.43626	-.39061	.40792	.06181	-.03871	-.05871	-.37349	.22527	-.33961	-.26883	-.06090	-.19646
(5) Arc SQRT Pretest	-.32329	-.27520	.40862	.15479	.03870	-.03191	-.37881	.16047	-.24895	-.21382	-.09874	-.12492
(6) Raw Posttest	-.44590	-.40234	.40834	.07635	-.04826	-.06422	-.34838	.21541	-.34416	-.27983	-.05432	-.18132
(7) Arc SQRT Posttest	-.01640	.02987	-.09399	-.03425	-.02212	-.04561	-.18734	-.08247	-.01074	.02705	-.01200	-.08043
(8) Class Size	.15197	.11790	-.23900	-.03425	-.02212	-.04561	.11463	-.12357	.15675	.19550	.10128	-.00285
(9) Class Absence	.14940	.24163	-.24148	-.15799	-.02208	-.00530	.03322	-.03929	.09417	.03927	.03194	.11322
(10) % Males	-.02160	-.11337	.00877	-.04221	-.02446	.08999	-.16343	.08756	-.02573	-.01655	.02525	.01748
(11) Class Turnover	-.09412	-.18086	.30934	.11890	.01656	-.08789	-.16797	.10281	-.10793	-.02422	-.07026	.07450
(12) Mean Material Level	.10110	.16418	-.19415	-.02073	.03190	.04253	-.01423	-.09438	.37806	.30866	.04373	.09800
(13) Variability Material Level	1.00000	.75730	-.41745	-.06187	-.02555	-.01045	.26245	-.14612	.25022	.29305	-.08158	.06991
(14) Reading on Own	.75730	1.00000	-.39492	-.07417	-.00010	-.00288	.11057	-.21227	.13936	.20912	-.05626	.13382
(15) Reading Interest	-.41745	-.39492	1.00000	.21234	.07615	-.02988	-.24341	.27296	-.25785	-.18345	-.24912	-.17131
(16) Need for Help	-.06187	.07417	.21234	1.00000	-.07281	-.00822	-.17306	.01890	-.10625	-.12037	-.14071	.00156
(17) Teacher Turnover	-.02555	-.00010	.07615	-.07281	1.00000	.14102	.05809	-.00772	-.12188	-.07122	.03011	-.10749
(18) Teacher Absence	.01045	.00288	-.02988	-.00822	.14102	1.00000	.02899	.10657	-.10320	-.01577	.06786	-.06188
(19) % Free Lunch	.26245	.11057	-.24341	-.17306	.05809	.02899	1.00000	-.31002	.20017	.19935	.22353	-.04699
(20) Math Specialist	-.14612	-.21227	.27296	.01890	-.00772	.10657	-.31002	1.00000	-.32056	-.35219	-.28385	-.01743
(21) Team Teaching	.25022	.13936	-.25785	-.10625	-.12188	-.10320	.20017	-.32056	1.00000	.62752	-.05553	.07157
(22) Open Space	.29305	.20912	-.18345	-.12037	-.07122	-.01577	.19935	-.35219	.62752	1.00000	.03933	-.01229
(23) Reading Problems	-.08159	-.05626	-.24912	-.14071	.03011	.06786	.22353	-.28385	-.05553	.03933	1.00000	.17007
(24) Behavior Barriers	.06991	.13352	-.17131	.00156	-.10749	-.06188	-.04699	-.01743	.07157	-.01229	.17007	1.00000
(25) Health Barriers	-.13119	-.09154	-.12785	-.11969	-.05583	-.04533	.04324	-.07325	.02307	-.01731	.28194	.13142

Table B p. 2.

		(25)
(1)	Grade	.04642
(2)	Test Level Count	.06463
(3)	Pretest Score	-.09395
(4)	Raw Pretest	-.01536
(5)	Arc SQRT Pretest	-.09393
(6)	Raw Posttest	-.01341
(7)	Arc SQRT Posttest	-.06925
(8)	Class Size	.02979
(9)	Class Absence	.07735
(10)	% Males	.04306
(11)	Class Turnover	.09766
(12)	Mean Material Level	.07690
(13)	Variability Material Level	-.13113
(14)	Reading on Own	-.09154
(15)	Reading Interest	-.12785
(16)	Need for Help	-.11969
(17)	Teacher Turnover	-.05593
(18)	Teacher Absence	-.04535
(19)	% Free Lunch	.04324
(20)	Math Specialist	-.07325
(21)	Team Teaching	.02307
(22)	Open Space	-.01731
(23)	Reading Problems	.23194
(24)	Behavior Barriers	.13142
(25)	Health Barriers	1.00000

APPENDIX F

Table C

Correlation Matrix for Intermediate Grades

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
(1) Grade	1.00000	.15294	.03187	-.08358	.03309	-.09094	-.00159	.01920	.01828	.02767	.57467	-.02703
(2) Test Level Count	.15294	1.00000	.15886	.18699	.15464	.19510	-.24276	-.04800	.03011	.29453	.01130	-.21352
(3) Pretest Score	.03187	.15886	1.00000	.75992	.99956	.75451	-.02399	-.13729	-.02283	.14717	.15911	-.19490
(4) Raw Pretest	-.08358	.18699	.75992	1.00000	.75849	.99554	-.00385	-.18418	-.05175	.19208	.07928	-.21962
(5) Arc SQRT Pretest	.03309	.15464	.99956	.75849	1.00000	.75438	-.02411	-.13788	-.04393	.14434	.16161	-.22626
(6) Raw Posttest	-.09094	.19510	.75451	.99554	.75438	1.00000	-.00973	-.19728	-.05546	.19209	.08116	-.21962
(7) Arc SQRT Posttest	-.00159	-.24276	-.02399	-.00385	-.02411	-.00973	1.00000	-.10954	-.04393	-.05175	-.16555	.20033
(8) Class Size	.01920	-.04800	-.13729	-.18418	-.13788	-.19728	-.10954	1.00000	.07795	.23246	-.09290	-.19465
(9) Class Absence	.01828	.03011	-.02283	-.05175	-.04393	-.05546	-.04393	.07795	1.00000	.01254	-.16943	-.12359
(10) % Males	.02767	.29453	.14717	.19208	.14434	.19209	-.05175	.23246	.01254	1.00000	-.10855	-.16546
(11) Class Turnover	.57467	.01130	.15911	.07928	.16161	.08116	-.16555	-.09290	-.16943	-.10855	1.00000	-.22626
(12) Mean Material Level	.02903	.21352	-.19490	-.21962	-.19465	-.22626	.20033	-.12532	-.25959	-.07124	-.25959	1.00000
(13) Variability Material Level	.05589	.09022	-.12491	-.15758	-.12359	-.16546	-.05029	-.24034	-.12531	-.13022	-.24059	.31635
(14) Reading on Own	.00794	.10041	-.20077	-.23665	-.20033	-.24059	.05174	-.27875	-.02716	-.16978	-.41778	.27352
(15) Reading Interest	.04073	-.06570	.20311	.26957	.20443	.27352	-.05174	-.12407	-.02007	-.11060	-.05800	-.01356
(16) Need for Help	.09678	-.08138	.11393	.10745	.11496	.10870	-.00503	-.10013	-.06940	-.07159	-.04137	.01823
(17) Teacher Turnover	-.04311	-.09053	.06694	.02142	.06900	.02409	-.07623	-.07263	-.04072	-.03252	-.05056	.04496
(18) Teacher Absence	-.03660	.10150	.01813	-.03105	.01986	-.02275	-.04994	-.02220	-.09660	-.01536	-.17157	.02578
(19) % Free Lunch	-.07759	-.03190	-.25609	-.26694	-.25775	-.26802	-.20220	-.28779	-.01595	-.17557	-.02578	-.18537
(20) Math Specialist	-.03256	-.02321	-.18953	-.14924	-.19003	-.14343	-.08440	-.14594	-.01536	-.16978	-.10750	-.14895
(21) Team Teaching	-.04256	-.03786	-.21425	-.20931	-.21259	-.20990	.16767	.07954	.00130	-.11060	-.05800	-.04137
(22) Open Space	.03809	-.06105	.00258	-.07560	.00274	-.07325	.05821	.07157	-.09807	-.03814	-.04137	-.01356
(23) Reading Problems	-.06963	-.00799	-.11531	-.13041	-.11489	-.13318	-.07678	.34337	.00928	-.10228	-.23074	.13333
(24) Behavior Barriers	-.12208	.00474	-.11735	-.13263	-.11697	-.13405	-.06405	.13270	.08563	-.03614	-.20201	.15113
(25) Health Barriers	-.12780	-.04976	.06452	-.00590	.06504	-.01007	-.05122	.07064	.04645	.02625	-.21947	.12065

	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)	(22)	(23)	(24)
(1) Grade	.05589	.00794	.04073	.09678	-.04311	-.03660	-.07759	-.03256	-.04256	.03809	-.06963	-.12208
(2) Test Level Count	.09022	.10041	-.06570	.08138	-.09053	.10150	-.03190	-.02321	-.03786	-.06105	-.00799	-.00474
(3) Pretest Score	-.12491	-.20077	.20311	.11393	.06694	.01813	-.25609	-.18953	-.21425	-.00258	-.11531	-.11735
(4) Raw Pretest	-.15758	-.23665	.26957	.10745	.02142	-.03105	-.26694	-.14924	-.20931	-.07560	-.13041	-.13263
(5) Arc SQRT Pretest	-.12359	-.20033	.20443	.11496	.06900	.01986	-.25775	-.19003	-.21259	-.00274	-.11489	-.11697
(6) Raw Posttest	-.16546	-.24059	.27352	.10870	.02409	-.02275	-.26802	-.14343	-.20990	-.07325	-.13318	-.13405
(7) Arc SQRT Posttest	.05029	.05174	-.05174	-.00503	-.07623	.04994	-.20220	-.08440	.16767	.05821	-.07678	-.06405
(8) Class Size	.25359	.27875	-.12407	-.10013	-.07263	-.04072	-.28779	-.14594	.07954	.07157	.34337	.13270
(9) Class Absence	.07174	.12031	-.02716	-.06940	-.04072	-.09660	-.01595	-.01536	.00130	-.09807	.00928	.08563
(10) % Males	-.06711	-.13002	.16939	.11060	-.07159	-.01536	-.17557	-.16978	-.11060	-.03814	-.10228	-.03614
(11) Class Turnover	-.20006	-.24656	.41778	.05800	-.04137	-.05056	-.33513	-.18537	-.14895	-.13335	-.23074	-.20201
(12) Mean Material Level	.21782	.31935	-.27645	-.01529	.01823	.04496	-.19295	-.09529	-.16148	-.15987	-.15505	-.15113
(13) Variability Material Level	1.00000	.74926	-.41332	-.00114	-.09699	.12285	-.24259	-.11719	-.10393	-.12139	-.06433	-.14333
(14) Reading on Own	.74926	1.00000	-.41586	-.00665	-.07691	.06103	-.25193	-.07659	-.13970	-.08228	-.09418	-.06415
(15) Reading Interest	-.41332	-.41586	1.00000	.02696	.05233	-.05546	-.34894	-.03940	-.10908	-.20455	-.00448	-.09015
(16) Need for Help	-.00114	-.00665	.02696	1.00000	.09035	-.12146	-.01254	-.05453	-.09692	-.10852	-.04794	-.06030
(17) Teacher Turnover	-.09699	-.07691	.05233	.09035	1.00000	-.07433	-.08605	-.04915	-.15449	-.01558	-.09436	-.03246
(18) Teacher Absence	.12285	.06103	-.05546	.12146	-.07433	1.00000	-.09324	-.08324	-.05287	-.09558	-.19550	-.12721
(19) % Free Lunch	.24259	.25193	-.34894	.01254	-.08605	-.09324	1.00000	-.25791	-.25791	-.20455	-.09436	-.09217
(20) Math Specialist	.09529	.11719	-.03940	.05453	-.04915	-.08324	-.25791	1.00000	-.04569	-.04158	-.04577	-.09017
(21) Team Teaching	.16149	.10393	-.13972	.10203	-.15449	-.05287	-.25791	-.04569	1.00000	-.05577	-.09577	-.05540
(22) Open Space	.15337	.12132	-.08229	.20455	-.01558	-.09558	-.20455	-.04158	-.05577	1.00000	-.24781	-.24781
(23) Reading Problems	.16329	.09648	-.22413	.04794	-.09436	-.19550	-.09436	-.04577	-.09577	-.24781	1.00000	-.22391
(24) Behavior Barriers	.12699	.14333	-.21447	.06030	-.03246	-.12721	-.09217	-.09017	-.05540	-.24781	-.22391	1.00000
(25) Health Barriers	.11055	.09614	-.21447	.05135	-.09423	.10428	.13909	-.04044	.14891	.10933	.22554	.22391

(25)

(1)	Grade	-.12780
(2)	Test Level Count	-.04876
(3)	Pretest Score	.06452
(4)	Raw Pretest	-.00580
(5)	Arc SQRT Pretest	.06504
(6)	Raw Posttest	-.01007
(7)	Arc SQRT Posttest	-.05122
(8)	Class Size	.07064
(9)	Class Absence	.04645
(10)	% Males	.02625
(11)	Class Turnover	-.21947
(12)	Mean Material Level	.12669
(13)	Variability Material Level	.16055
(14)	Reading on Own	.08614
(15)	Reading Interest	-.21647
(16)	Need for Help	-.05135
(17)	Teacher Turnover	-.09423
(18)	Teacher Absence	.10429
(19)	% Free Lunch	.13909
(20)	Math Specialist	-.04044
(21)	Team Teaching	.14891
(22)	Open Space	.10933
(23)	Reading Problems	.22559
(24)	Behavior Barriers	.39391
(25)	Health Barriers	.00000

Table D

APPENDIX G

Number of Students and Classes in 34 Evaluation System Schools by Year and Grade

School Year	Official October Membership	Total Number of Classes	Number of Classes in the Analysis Pool (1)	Usable Sample for Analysis (2)			Usable Sample by Grade Level (3)						
				Primary (1-3)	Inter-Mediate (4-6)	Total	Individual Grades 1	2	3	4	5	6	Total
1975-76	16,170	627	512	108	121	229	71	55	71	70	74	59	400
1976-77	15,382	570	551	191	205	396	57	62	53	65	63	76	376

(1) Classes wherein the average membership of students was less than 100 days were excluded from the analysis pool. Additionally, for a class to be included certain ID matching criteria had to be met.

(2) A class was not included in the usable sample for analysis if a student within the class was missing data on any one of the selected variables. Only students with pre-post matched data on the same level PRT were used. Combinations, or multi-level grade classes were included.

(3) Classes composed of combinations, or multi-level grades were excluded from the by grade level analyses.

202